

Multilevel Datenfusion konkurrierender Sensoren in der Fahrzeugumfelderfassung

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach Informatik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät II

Humboldt-Universität zu Berlin

von

Dipl.-Inf. Mathias Haberjahn

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:

Prof. Dr. Elmar Kulke

Gutachter:

1. Prof. Dr. rer. nat. Ralf Reulke
2. Prof. Dr.-Ing. Christoph Stiller
3. Prof. Dr.-Ing. Peter Eisert

eingereicht am: 28.03.2013

Tag der mündlichen Prüfung: 29.10.2013

Abstract

With the present thesis a contribution to the increase of the accuracy and reliability of a sensor-supported recognition and tracking of objects in a vehicle's surroundings should be made. Based on a detection system, consisting of a stereo camera and a laser scanner, novel developed procedures are introduced for the whole processing chain of the sensor data. In addition, a new framework is introduced for the fusion of heterogeneous sensor data. By combining the data fusion results from the different processing levels the object detection can be improved.

After a short description of the used sensor setup the developed procedures for the calibration and mutual orientation are introduced. As a result, the measuring error for the laser scanner can be reduced by over 10 %.

With the segmentation of the spatial point data existing procedures are extended by the inclusion of measuring accuracy and specificity of the sensor. Furthermore, an approach for the determination of the minimally enclosing object box and the object alignment to support the object tracking is introduced.

In the subsequent object tracking a new computation-optimized approach for the association of the related object hypotheses is presented. In addition, a model for a dynamic determination and tracking of an object reference point is described which exceeds the classical tracking of the object center in the track accuracy.

By the introduced fusion framework it is possible to merge the sensor data at three different processing levels (point, object and track level). A sensor independent approach for the low fusion of point data is demonstrated which delivers the most precise object description in comparison to the other fusion levels and the single sensors. For the higher fusion levels new procedures were developed to discover and clean up the detection and processing mistakes benefiting from the competing sensor information. Finally it is described how the fusion results of the upper and lower levels can be brought together for an ideal object description. The effectiveness of the newly developed methods was checked either by simulation or in real measurement scenarios.

The interdisciplinary standard of this thesis lies in the novel concept of a combined multi layer fusion of heterogeneous sensor data for a more reliable and more precise object description in the vehicle surrounding field. This concept was implemented in his totality with a technological innovative reference sensor system and tested successfully.

Zusammenfassung

Mit der vorliegenden Dissertation soll ein Beitrag zur Steigerung der Genauigkeit und Zuverlässigkeit einer sensorgestützten Objekterkennung und -verfolgung im Umfeld eines Fahrzeuges geleistet werden. Aufbauend auf einem Erfassungssystem, bestehend aus einer Stereokamera und einem Mehrzeilen-Laserscanner, werden teils neu entwickelte Verfahren für die gesamte Verarbeitungskette der Sensordaten vorgestellt. Zusätzlich wird ein neuartiges Framework zur Fusion heterogener Sensordaten eingeführt, welches über eine Zusammenführung der Fusionsergebnisse aus den unterschiedlichen Verarbeitungsebenen in der Lage ist, die Objektbestimmung zu verbessern.

Nach einer kurzen Beschreibung des verwendeten Sensoraufbaus werden die entwickelten Verfahren zur Kalibrierung und gegenseitigen Orientierung des Sensorpaares vorgestellt. Durch das eingeführte Fehlermodell kann hierbei der Messfehler beim Laserscanner um über 10 % reduziert werden.

Bei der Segmentierung der räumlichen Punktdaten werden bestehende Verfahren durch die Einbeziehung von Messgenauigkeit und Messspezifik des Sensors erweitert. Ferner wird ein Ansatz zur Bestimmung der minimal umschließenden Objekthülle und der Objektausrichtung zur Stützung der Objektverfolgung vorgestellt.

In der anschließenden Objektverfolgung wird neben einem neuartigen berechnungsoptimierten Ansatz zur Assoziation der Objekthypothesen ein Modell zur dynamischen Referenzpunktbestimmung und -Verfolgung beschrieben, das die klassische Verfolgung der Objektmitte in der Trackgenauigkeit übertrifft.

Durch das vorgestellte Fusions-Framework ist es möglich, die Sensordaten wahlweise auf drei unterschiedlichen Verarbeitungsebenen (Punkt-, Objekt- und Track-Ebene) zu vereinen. Hierzu wird ein sensorunabhängiger Ansatz zur Fusion der Punktdaten dargelegt, der im Vergleich zu den anderen Fusionsebenen und den Einzelsensoren die genaueste Objektbeschreibung liefert. Für die oberen Fusionsebenen wurden unter Ausnutzung der konkurrierenden Sensorinformationen neuartige Verfahren zur Bestimmung und Bereinigung der Detektions- und Verarbeitungsfehler entwickelt. Abschließend wird beschrieben, wie die fehlerreduzierenden Verfahren der oberen Fusionsebenen mit der optimalen Objektbeschreibung der unteren Fusionsebene für eine optimale Objektbestimmung zusammengeführt werden können. Die Effektivität der neu entwickelten Verfahren wurde entweder durch Simulation oder in realen Messszenarien überprüft.

Der Anspruch der Arbeit besteht in dem neuartigen Konzept zur kombinierten Mehrebenenfusion heterogener Sensordaten für eine zuverlässigere und genauere Objektbestimmung im Fahrzeugumfeld. Das Konzept wurde in seiner Gesamtheit mit einem technologisch innovativen Referenzsensordaten-System umgesetzt und erfolgreich getestet.

Inhaltsverzeichnis

1	Einleitung	1
2	Fahrzeugumfelderfassung	7
2.1	Sensordatenfusion	7
2.1.1	Übersicht der Fusionsarchitektur	8
2.1.2	Übersicht der Prozess-Modelle	10
2.2	Sensorsysteme	11
2.2.1	Erfassung mit homogenen Sensoren	11
2.2.2	Erfassung mit heterogenen Sensoren	12
2.3	Konzepte zur Sensorfusion	14
2.4	Multilevel-Multisensor-Fusion	14
2.5	Zusammenfassung	18
3	Sensorsystem	19
3.1	Aufbau des Sensorsystems	19
3.1.1	Spezifikation des Stereosystems	19
3.1.2	Spezifikation des Laserscanners	23
3.2	Kalibrierung der Sensoren	24
3.2.1	Kalibrierung der Stereokamera	24
3.2.2	Kalibrierung des Laserscanners	27
3.2.3	Bestimmung der Sensorsichtbereiche	37
3.3	Ableitung der Datengenauigkeit	37
3.3.1	Genauigkeit der Stereodaten	38
3.3.2	Genauigkeit der Laserdaten	38
3.4	Filterung der Punktdaten	39
3.4.1	Bestimmung des Untergrundes	40
3.4.2	Reduktion der Stereodaten	40
3.5	Zusammenfassung	42
4	Mid Level - Segmentierung und Objektbestimmung	43
4.1	Segmentierung	43
4.1.1	Segmentierungsmodell	44
4.1.2	Segmentierung der Punktinformationen	46
4.1.3	Segmentierungswahrscheinlichkeit	46
4.2	Segmentverfeinerung	49
4.2.1	Zusammenfassung von Segmenten	49

4.2.2	Bestimmung der Objektausrichtung	49
4.2.3	Ableitung der Objektbox	53
4.2.4	Segmentierungsfehler	56
4.3	Zusammenfassung	58
5	High Level - Objektverfolgung	59
5.1	Objektassoziiierung	60
5.1.1	Metrik der Objektdistanzen	63
5.2	Objekttracking	66
5.2.1	Prädiktion des Objektzustandes	67
5.2.2	Korrektur des Objektzustandes	69
5.2.3	Fehler in der Objektverfolgung	70
5.3	Objektbeobachtbarkeit	71
5.3.1	Referenzpunktbestimmung	71
5.3.2	Referenzpunktwechsel	74
5.4	Zusammenfassung	77
6	Multilevel-Datenfusion	79
6.1	Fusion der Punktdaten	80
6.2	Fusion der Objektdaten	82
6.2.1	Assoziierung der Multisensor-Objekte	82
6.2.2	Auflösung von Fragmentierung und Zusammenfassung	84
6.2.3	Auflösung von Fehl- und Falschdetektionen	88
6.2.4	Zustandsfusion	91
6.3	Fusion der Tracks	93
6.3.1	Assoziierung der Tracks	93
6.3.2	Auflösung von Fehl- und Falschdetektionen	95
6.3.3	Auflösung von Multi-Objekt-Tracks	97
6.3.4	Zustandsfusion	99
6.4	Kombination der Fusionsebenen	101
6.5	Zusammenfassung	103
7	Experimentelle Ergebnisse	105
7.1	Messgenauigkeiten der Sensoren	105
7.2	Tracking mit Referenzpunktwechsel	107
7.3	Auflösung von Falsch- und Fehldetektion auf Objektebene	116
7.4	Auflösung der Fragmentierung und Zusammenfassung von Objekten	120
7.5	Auflösung von Multi-Objekt-Tracks	126
7.6	Genauigkeit der Objektbestimmung	130
7.6.1	Vergleich der Sensoren	132
7.6.2	Vergleich der Fusionsebenen	139
7.7	Zusammenfassung	151
8	Zusammenfassung und Ausblick	153

Anhang A	167
1 Transformation der Referenzpunkte	167
2 Parameter zur Experimentauswertung	168

Abkürzungen

Abkürzung	Erklärung
ANA	engl. <i>All Neighbours Association</i>
BKS	Bildkoordinatensystem
BM	engl. <i>Block Matcher</i>
DDF	engl. <i>Deferred Decision Logic</i>
EKF	Erweitertes Kalman Filter
EOT	Einzel-Objekt-Track
FCIF	engl. <i>Fast Covariance Intersection Fusion</i>
FNO	Fehl-Objekt engl. <i>False Negative Object</i>
FNT	Fehl-Track engl. <i>False Negative Track</i>
FNZ	Fehlzuordnung
FO	Fragmentierungs-Objekt
FPO	Falsch-Objekt engl. <i>False Positive Object</i>
FPT	Falsch-Track engl. <i>False Positive Track</i>
FPZ	Falschzuordnung
GNN	engl. <i>Global Nearest Neighbour</i>
IF	engl. <i>Information Fusion</i>
ILP	engl. <i>Independent Likelihood Pool</i>
IMM	engl. <i>Interacting Multiple Model</i>
IOP	engl. <i>Independent Opinion Pool</i>
KTC	Konflikt-Track-Cluster
LKS	Laserscanner-Koordinatensystem
MHT	engl. <i>Multi Hypotheses Tracking</i>
MMFF	Multilevel Multisensor Fusion Framework
MOT	Multi-Objekt-Track
MPT	Mittelpunkt-Tracking
MTT	engl. <i>Multi Target Tracking</i>
OBF	Objektboxfehler
OKS	Objektkoordinatensystem
OZF	Objektzustandsfehler
PDA	engl. (probabilistic data association)
RMSE	engl. <i>Root Mean Squared Error</i>
RPT	Referenzpunkt-Tracking
SCF	engl. <i>Simple Convex Fusion</i>
SDL	engl. <i>Sequential Decision Logic</i>
SKKS	Stereokamera-Koordinatensystem

Inhaltsverzeichnis

SKS	Sensorkoordinatensystem
STT	engl. <i>Single Target Tracking</i>
SVD	Singulär-Wert-Zerlegung engl. <i>Singular Value Decomposition</i>
SW	Segmentierungswahrscheinlichkeit
TLS	Terrestrischer Laserscanner
UNA	engl. <i>Unique Neighbour Association</i>
WLS	engl. <i>Weighted Least Squares</i>
ZO	Zusammenfassungs-Objekt

Symbolverzeichnis

Symbol	Erklärung
A	Designmatrix
B	Basislänge zwischen den Projektionszentren der Stereokameras
c	Kammerakonstante, Brennweite
C	Mittelpunkt der Objektbox
d	Disparität im Stereo-Normalfall
d_x	Objektausmaß in X-Richtung des OKS
d_y	Objektausmaß in Y-Richtung des OKS
d_z	Objektausmaß in Z-Richtung des OKS
$d_M^{p_1}(p_1, p_2)$	Mahalanobis-Distanzfunktion von p_1 zu p_2
$d_{\bar{M}}(p_1, p_2)$	Mahalanobis-Distanzfunktion mit gemittelter Kovarianz der Zufallsgrößen p_1 und p_2
d_{F-}	Trennungsdistanz zwischen zwei Objekthypothesen eines Sensors
d_{F+}	Überlappungsdistanz zwischen einer Objekthypothese und einem Cluster
G	Assoziationsmatrix
H	Matrix zur Transformation von Messwert- in Zustandsvektor
K	Rotationszentrum
K	Matrix zum Kalman-Gain
O	Projektionszentrum der Kamera
$P(S_P^X)$	Segmentierungswahrscheinlichkeit am Punkt P des Sensors X
P^-	A priori Kovarianzmatrix von \hat{x}^-
P	A posteriori Kovarianzmatrix von \hat{x}
Q	Matrix zum Prozessrauschen des Bewegungsmodells
R	Rotationsmatrix; Matrix zum Messrauschen
R_x, R_y, R_z	Rotationsmatrizen zur Rotation an den Koordinatenachsen
R_L	Rotationsmatrix zur Transformation zwischen LKS und SKS
r	Abstand eines Punktes vom Koordinatenursprung
T	Translationsvektor
v	Objektgeschwindigkeit
V_R	Volumen des Referenzobjektes
V_U	Überlappungsvolumen zwischen Mess- und Referenzobjekt
\hat{X}^-	Geschätzter a priori Objektzustandsvektor
\hat{X}	Geschätzter a posteriori Objektzustandsvektor
\hat{Z}	Geschätzter Messwertvektor
ϵ	Beliebig kleine Zahl größer als Null

Inhaltsverzeichnis

θ	Winkel zwischen positiver X-Achse und dem Vektor \vec{p} in der X-Z-Ebene des LKS
κ	Rotationswinkel der Z-Achse (photogrammetrische Orientierung)
μ	Erwartungswert einer Zufallsgröße
ν	Zufallsgröße des Prozessrauschens
ξ	Zufallsgröße des Messrauschens
σ	Standardabweichung einer Zufallsgröße
σ^2	Varianz einer Zufallsgröße
ϕ	Rotationswinkel der Y-Achse (photogrammetrische Orientierung)
φ	Winkel zwischen positiver X-Achse und dem Vektor \vec{p} in der X-Y-Ebene des LKS
ψ	Winkel zwischen Richtungsvektor des Objekts und der X-Achse in der X-Y-Ebene des SKS
ω	Rotationswinkel der X-Achse (photogrammetrische Orientierung); Winkelgeschwindigkeit
\mathcal{F}^P	Fehler-Ellipsoid für einen Punkt P
\mathcal{R}_O^O	Objektraum eines Objektes O
\mathcal{S}^P	Segmentierungs-Ellipsoid für einen Punkt P
\mathcal{R}_V^O	Verdeckungsraum eines Objektes O
\mathbb{D}	Menge an Sensoren, mit Objekt/Track-Detektion
$\bar{\mathbb{D}}$	Menge an Sensoren, ohne Objekt/Track-Detektion
\mathbb{U}	Menge an Sensoren, mit Objektdetektion über mehrere Cluster
$\bar{\mathbb{U}}$	Menge an Sensoren, mit Objektdetektion über genau ein Cluster

1 Einleitung

Die absoluten Unfälle mit schweren Personenschäden oder tödlichem Ausgang sind in den vergangenen Jahren konstant rückläufig und haben heute in Deutschland den Tiefstwert seit 60 Jahren erreicht. Trotzdem sterben nach Angaben des Statistischen Bundesamtes [104] durchschnittlich täglich zehn Menschen und mehr als 1000 werden verletzt und davon über 170 schwer (siehe Abbildung 1.1).

Neben den Personenschäden wird zudem der volkswirtschaftliche Schaden beispielsweise für das Jahr 2008 auf ca. 31 Milliarden Euro geschätzt, wobei die Kosten annähernd zu gleichen Teilen aus Sach- und Personenschäden hervorgehen [105].

Laut Statistischem Bundesamt wurden auch im Jahr 2010 84 % aller Unfälle durch ein Fehlverhalten des Fahrzeugführers verursacht. Allgemeine Ursachen (Witterung, Straßenverhältnisse oder Hindernisse) sowie ein Fehlverhalten anderer Verkehrsteilnehmer hatten hierbei zusammen nur einen Anteil von ca. 15.7 % (nach [103]).

Wie aus Abbildung 1.2 zu entnehmen, ist eine nicht angepasste Geschwindigkeit die Hauptursache für vom Fahrer verursachte Unfälle mit Personenschäden. Weitere Ursachen liegen beim fehlerhaften Abbiegen, Wenden, Rückwärtsfahren sowie Ein- und Anfahren [104].

Die *100 Car Study* besagt, dass in fast 80 % aller Unfälle und ca. 65 % aller Beinahe-Unfälle die Unachtsamkeit des Fahrers im kurzen Zeitraum vor dem Vorfall (z.B. 3 s) eine entscheidende Teilursache darstellt [28]. Für 2/3 aller Unfälle in denen die Unachtsamkeit des Fahrers eine Rolle spielt, wurden die Ursachen folgendermaßen aufgeschlüsselt:

- 46.7 % Objekte außerhalb des Fahrzeugs
- 22.5 % Bedienung von Radio/CD
- 17.3 % Mitfahrer
- 6.8 % Bewegende Objekte im Fahrzeug
- 4.1 % Essen, Trinken, Rauchen
- 2.4 % Telefonieren

1 Einleitung

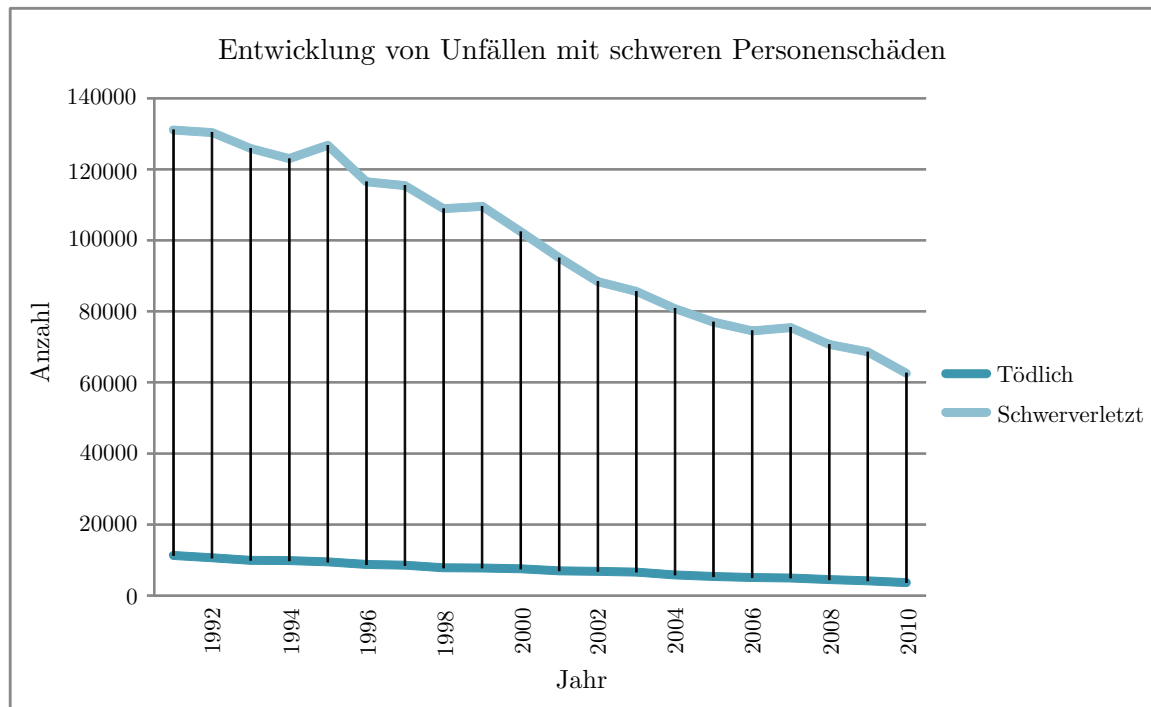


Abbildung 1.1: Entwicklung der Unfälle mit schweren Personenschäden oder tödlichem Ausgang (nach [103])

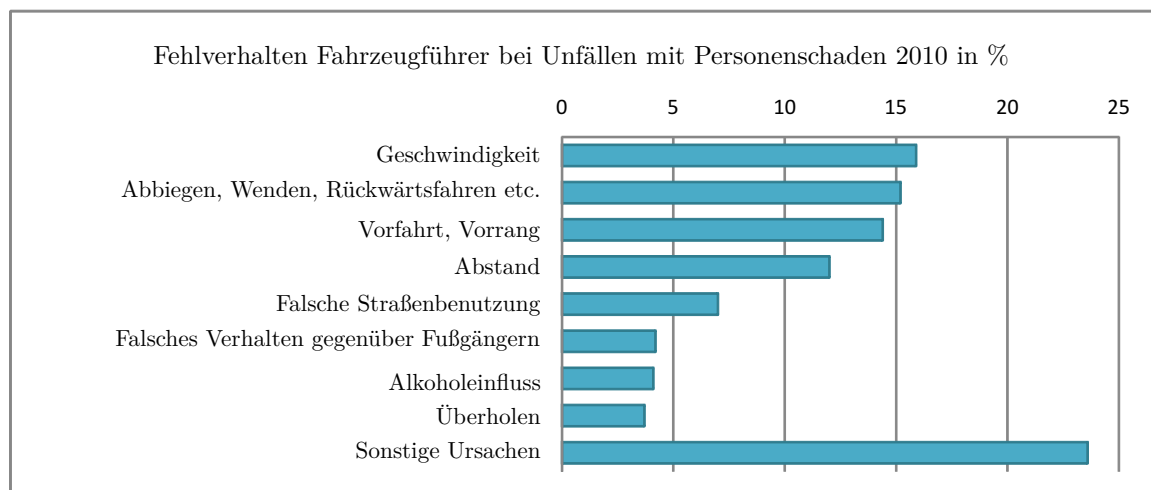


Abbildung 1.2: Fehlverhalten des Fahrzeugführers bei Unfällen mit Personenschäden (nach [104])

Die Unachtsamkeit des Fahrers lässt sich einerseits auf eine Unterforderung in der aktuellen Fahrsituation zurückführen. Eine weitere Hauptursache für ein fehlerhaftes Fahrverhalten stellt die Überforderung des Fahrers dar. Abbildung 1.3 zeigt die schon im Jahr 1908 erkannte Abhängigkeit zwischen der Qualität der menschlichen Fahrleistung und dem Schwierigkeitsgrad der aktuellen Fahrsituation [115]. Demnach erscheint eine Unterstützung durch eine Fahrassistenz sowohl in der fahrerseitigen Unterforderung (z.B. Nachtfahrt auf Landstraße) als auch in der Überforderung (z.B. dichter Verkehr auf Autobahn) als sinnvoll, da hier die Performanz in der Fahrleistung vergleichsweise am geringsten ist.

Ein Lösungsansatz, um den entscheidenden Unsicherheitsfaktor Mensch bei der Erkennung und Vermeidung von gefährlichen Situationen zu unterstützen, ist die rechnergestützte Fahrassistenz (*Advanced Driver Assistance Systems* – ADAS).

Die sensorgestützte Fahrzeugumfelderfassung stellt hierbei die Grundlage der automatisierten Fahrassistenz dar. Sie liefert die Informationen über das aktuelle Verkehrsgeschehen im Umfeld des Fahrzeugs. Hierzu zählen z.B. die Lage, Größe und Bewegung von statischen und dynamischen Objekten bzw. hieraus abgeleitete Informationen zur Objektklasse und den wechselseitigen Interaktionen. Diese Informationen bilden die Entscheidungsgrundlage für eine nötige Assistenz. Für eine zuverlässige Assistenz müssen demnach fehlerhafte und fehlende Informationen auf ein Minimum reduziert und die Güte der Objektbestimmung maximiert werden.

Um die unvermeidbaren verfahrensspezifischen Nachteile homogener Sensornetze zu vermeiden, werden verstärkt heterogene Sensoren mit überlappenden Messbereichen in der Umfelderkennung erforscht und eingesetzt.

Für einen zuverlässigen Einsatz muss die Fehl- und Falschalarmrate jedoch weiter reduziert bzw. die Genauigkeit der Objekt- und Situationsbeschreibung gesteigert werden. Die Ursachen der Fehler liegen in der Sensorspezifik (z.B. negative Umwelteinflüsse auf das Messverhalten, Messgenauigkeit), der Datenverarbeitung (z.B. Fragmentierung und Zusammenfassung von Objekten, Assoziationsfehler, unflexibles Dynamikmodell im Tracking) oder den Messbedingungen (z.B. Verdeckung).

An diesem Punkt knüpft die Arbeit an und versucht über eine konkurrierende Fusion heterogener Sensordaten auf unterschiedlichen Verarbeitungsebenen die Fahrzeugumfelderfassung in ihrer Zuverlässigkeit und Genauigkeit zu steigern. Ein hierfür entwickeltes Fusions-Framework erlaubt die parallele Zusammenführung der Sensordaten auf den Verarbeitungsebenen der räumlichen Punktinformationen (*Low-Level*), der Objektdaten (*Mid-Level*) und der Track-Daten (*High-Level*). Die durch einen gemeinsamen Erfassungsbereich resultierende Informationsredundanz wird hierbei genutzt, um fehlerhafte und fehlende Detektionen oder Verarbeitungsfehler einzelner Sensoren anhand der widersprüchlichen Sensoraussagen zu erkennen. Die erkannten Konflikte sollen anschließend in einem konkurrierenden Vergleich der lokalen Sensoreigenschaften, beispielsweise der Messgenauigkeit oder der Segmentierungszuverlässigkeit, auf den oberen Fusionsebenen aufgelöst werden. Es wird weiter angenommen, dass die Datenfusion auf der untersten Ebene, aufgrund der höchsten Informationsdichte, die vergleichsweise höchste Genauigkeit in der Bestimmung der Objekthypothesen liefert. Diese beiden Aspekte sollen in einem kombinierten Verfahren zusammengeführt werden.

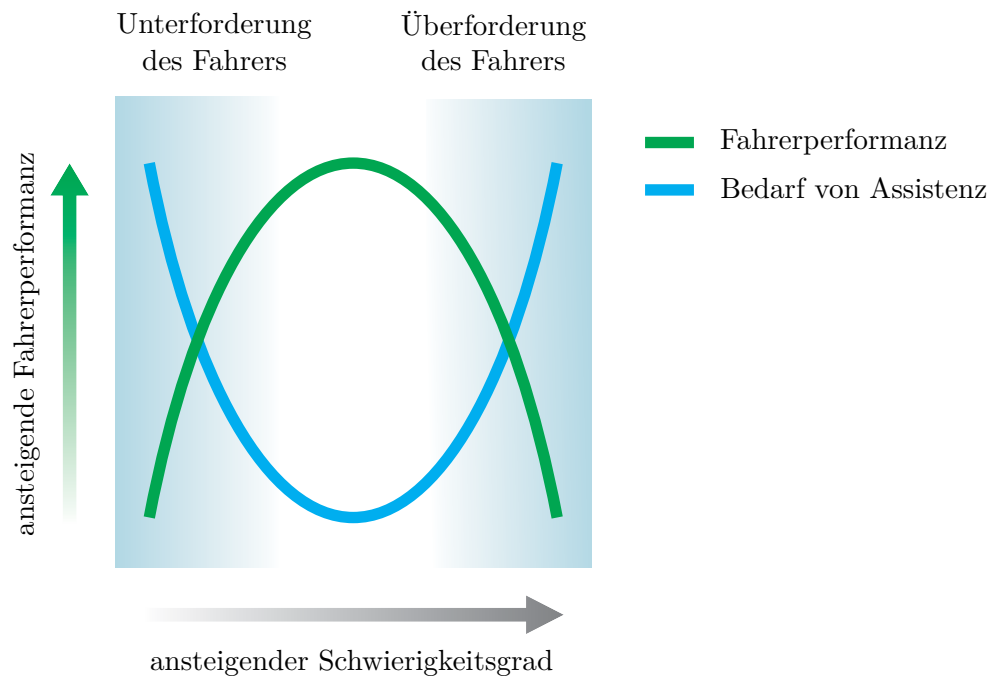


Abbildung 1.3: Abhängigkeit der Qualität der Fahrleistung vom Schwierigkeitsgrad der Fahr-situation nach [115]

Die Kernthesen dieser Arbeit lassen sich somit zusammenfassen zu:

- **Optimale Zustandsbestimmung der Objekte durch Low-Level-Fusion**
- **Erkennung und Entfernung von Mess- und Prozessierungsfehlern durch Fusion auf höheren Verarbeitungsebenen**
- **Reduzierung von Prozessierungsfehlern bei gleichzeitig optimaler Objektbestimmung durch eine kombinierte Fusion auf verschiedenen Prozessierungsebenen**

Eine Einführung in die übergeordnete Thematik der sensorgestützten Umfelderkennung aus dem Fahrzeug sowie der Sensordatenfusion liefert Kapitel 2. Es folgt eine diesbezügliche Einordnung der Arbeit in den aktuellen Stand der Technik und die Vorstellung des neu entwickelten Konzeptes zur Multilevel-Multisensor Fusion.

Im Kapitel 3 wird der Aufbau und die Funktionsweise des entwickelten Referenzsensorsystems, bestehend aus einem Mehrzeilen-Laserscanner und einer Stereokamera, vorgestellt. Hierbei implementierte neuartige Verfahren zur Kalibrierung und Registrierung der Einzelsensoren und des Sensorsystems sowie zur Vorverarbeitung der Rohdaten werden im Detail beschrieben.

Die Verfahren zur Prozessierung der Sensordaten auf der mittleren und oberen Ebene werden in den Kapiteln 4 und 5 erläutert. Hierzu zählen die Segmentierung der Punktinformationen,

die anschließende Bestimmung der Objekthypothesen sowie die Ableitung des Bewegungsverlaufs der Objekte im Tracking.

Die unterschiedlichen Verfahren zur Datenfusion auf den drei Ebenen sowie ein neuartiger Ansatz zur Kombination der Fusionsergebnisse zur Steigerung von Genauigkeit und Zuverlässigkeit der Umfelderkennung werden in Kapitel 6 dargestellt. Die experimentelle Auswertung und Validierung der vorgestellten Verfahren werden in Kapitel 7 vorgenommen.

Abschließend werden die Ergebnisse in Hinblick auf die Forschungsfragen der Arbeit in Kapitel 8 zusammengefasst und ein Ausblick auf mögliche Anknüpfungspunkte für zukünftige Arbeiten gegeben.

2 Fahrzeugumfelderfassung

Wie bereits in der Einleitung erwähnt, bildet die sensorgestützte Fahrzeugumfelderfassung die Basis der rechnergesteuerten Fahrerassistenz. Für eine umfassende Erfassung und zur Erhöhung der Zuverlässigkeit werden zumeist mehrere Sensoren eingesetzt. Zur Ableitung einer einheitlichen Beschreibung der Umwelt werden die verschiedenen Datenquellen in der Datenfusion zusammengefasst. Im Folgenden wird in den Themenbereich der Datenfusion, insbesondere in Hinblick auf die Fahrzeugumfelderfassung, eingeführt und der diesbezügliche aktuelle Stand der Technik dargestellt. Darauf aufbauend wird das entwickelte Framework zur Multilevel-Multisensor-Fusion und die Neuerung zum Stand der Technik vorgestellt.

2.1 Sensordatenfusion

Mit Sensor- oder Multisensor-Datenfusion werden im Allgemeinen die Methoden zur Kombination von Sensordaten bezeichnet, die zu einer Aufwertung der Information führen, die ohne die Datenzusammenführung nicht möglich wäre. Hierbei müssen die Daten nicht zwangsläufig von verschiedenen unabhängigen Sensorquellen stammen. So können auch zeitlich versetzte Messwerte eines Sensors für eine genauere Gesamtaussage kombiniert werden.

Wie für viele Technologien lassen sich auch für die Datenfusion Analogien in der Natur finden. So verfügen Menschen und Tiere über mehrere Sinne, die in ihrer Gesamtheit erst eine umfassende Umweltwahrnehmung ermöglichen und somit das Überleben sichern.

In [77] werden die Vorteile der Datenfusion in vier Punkten zusammengefasst:

Repräsentanz Die aus der Fusion resultierende Information besitzt eine höhere Abstraktionsebene oder Granularität als die einzelnen Datenquellen.

Bestimmtheit Durch die Fusion V_F wird eine Verstärkung der Aussagekraft oder Glaubwürdigkeit $p(V)$ eines Einzelsensors V mit $p(V_F) > p(V)$ erwartet.

Genauigkeit Die Standardabweichung der Daten ist nach der Fusion kleiner als für die eingehenden Einzelsensoren.

Vollständigkeit Jeder neue Informationszusatz trägt zur Komplettierung der Umfeldwahrnehmung bei.

Eine Unterteilung in quantitative und qualitative Vorzüge der Datenfusion wird in [54] vorgenommen. Die quantitativen Vorteile stellen den Genauigkeitsgewinn für die zu beobachtenden Größen dar (z.B. Positionsbestimmung, Objektklassifikation etc.).

2 Fahrzeugumfelderfassung

Die qualitativen Vorteile werden hiernach zusammengefasst mit:

- Erhöhung der Ausfallsicherheit
- Erweiterung der räumlichen Abdeckung
- Erweiterung der zeitlichen Abdeckung
- Erhöhung der Detektionssicherheit
- Verminderung von Mehrdeutigkeiten
- Erhöhung der räumlichen Messauflösung

In Abhängigkeit zu den eingesetzten Sensorverfahren und der räumlichen Ausrichtung der Sensoren wird häufig in drei verschiedene Fusionsmodi unterschieden [77]. So ist von einer **komplementären Fusion** (*complementary*) zu sprechen, wenn die Sensoren nicht direkt von einander abhängig sind, aber zusammen genommen eine komplettere Beschreibung des zu beobachtenden Phänomens liefern. So stellen beispielsweise zwei Radarsensoren, welche in Fahrtrichtung und entgegengesetzt am Fahrzeug angebracht sind, eine komplementäre Sensorkonfiguration dar.

Bei einer **konkurrierenden Fusion** (*competitive*) ist jeder Sensor in der Lage eine bestimmte Messgröße zu erfassen. Durch die redundante Mehrfachbeobachtung der Messgröße lässt sich eine genauere Gesamtmessung ableiten und fehlerhafte Messungen vermeiden. Beispielsweise kann durch eine Kopplung mehrerer Inertialmesssysteme zur Bestimmung für die translatorischen und rotatorischen Beschleunigungen eines Objekts der Gesamtfehler verringert werden.

In der **kooperativen Fusion** (*cooperative*) wird aus der Verknüpfung der Sensordaten eine Gesamtinformation abgeleitet, die von den einzelnen Sensoren nicht messbar ist. So kann ein aus zwei Kameras bestehendes Stereokamerasystem räumliche Informationen messen, wohingegen die einzelnen Kameras nur bildhafte Informationen liefern.

2.1.1 Übersicht der Fusionsarchitektur

Grundsätzlich können die Daten mehrerer Quellen in einer Fusionsarchitektur entweder an einem zentralen oder an mehreren dezentralen Fusionsknoten zusammengeführt werden. Als dritte Möglichkeit besteht als hybride Form die hierarchische Fusion (siehe Abbildung 2.1).

Die zentrale Fusion an einem Knoten stellt theoretisch die optimale Form der Datenfusion dar, da alle verfügbaren Daten in einem Schritt zusammengeführt werden. Hierfür müssen die Daten in einem gemeinsamen Format vorliegen. Die durch die Zentralität bedingten Nachteile liegen hingegen in (nach [77]):

Kommunikation Die gemeinsame Kommunikation zum zentralen Knoten kann einen Engpass darstellen.

Inflexibilität Die zentrale Fusion ist zumeist auf eine spezifische Anwendung ausgelegt, wodurch die Übertragung auf eine andere Anwendung oft einen erhöhten Mehraufwand bedeutet.

Verwundbarkeit Bei einem Ausfall des zentralen Knotens ist der gesamte Prozess unterbrochen.

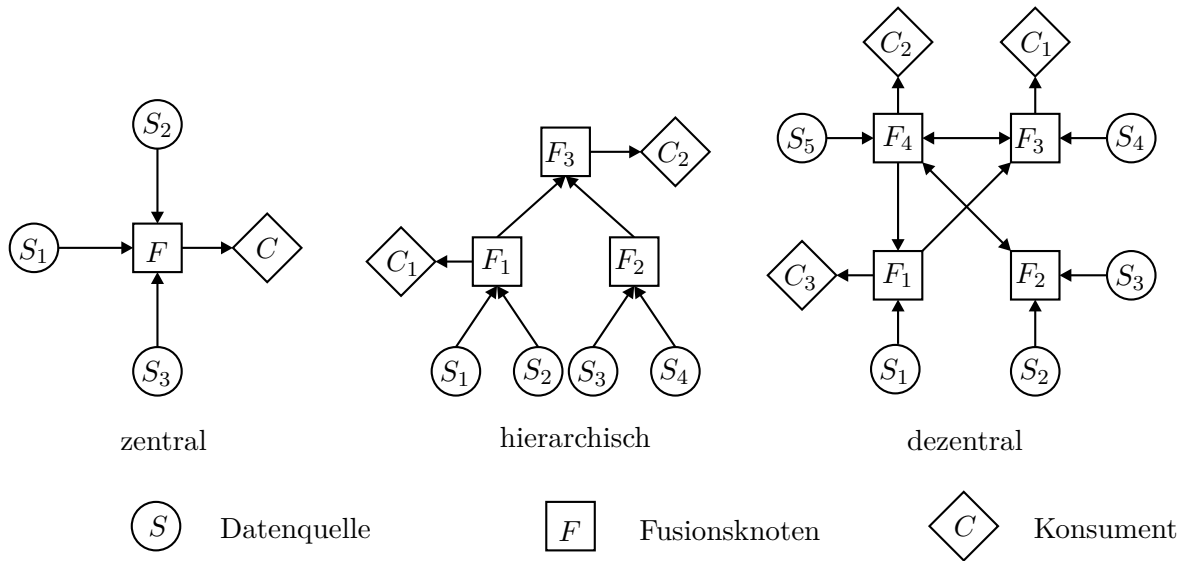


Abbildung 2.1: Architekturformen der Datenfusion (zentral, dezentral und hierarchisch) nach [69]

Nicht-Modularität Eine Erweiterung durch zusätzliche Datenquellen hat oft eine Anpassung des gesamten Fusionsknotens zur Folge.

Bei einer dezentralen oder verteilten Fusionsarchitektur werden die Daten an lokalen Fusionsknoten zusammengeführt. Dabei ist ein Fusionsknoten typischerweise nur mit einer Untermenge aller Datenquellen direkt verbunden. Des Weiteren fungieren die Fusionsknoten im Netz ebenfalls als Datenquelle. Somit bilden die Nachteile des zentralen die Vorteile des dezentralen Ansatzes:

Kommunikation Die Kommunikationslast verteilt sich auf mehrere Knoten.

Skalierbarkeit Aufgrund der fehlenden Einschränkung durch eine zentrale Rechen- oder Kommunikationseinheit ist das dezentrale System besser zu skalieren.

Robustheit Der Verlust von Sensorquellen oder Fusionsknoten muss nicht zwangsläufig die Gesamtfunktionalität beeinträchtigen.

Modularität Der dezentrale Ansatz unterstützt den modularen Aufbau der Fusionsarchitektur.

Der Nachteil der Dezentralität liegt in der Datenredundanz durch wiederholt fusionierte Information. Durch diese mehrfach auftretende gemeinsame Information an verschiedenen Knoten kann es in der Fusion zu einer verzerrten Gewichtung kommen, wodurch die Resultate verfälscht werden.

Die hierarchische Fusionstopologie versucht die Vorteile der beiden Ansätze zu kombinieren. Die Fusionsknoten sind hierarchisch in aufsteigenden Ebenen angeordnet, wobei der Datenfluss von den unteren zu den oberen Ebenen verläuft. Nach dem Durchlaufen von mehreren Fusions Ebenen werden die Daten am oberen zentralen Fusionsknoten zusammengeführt.

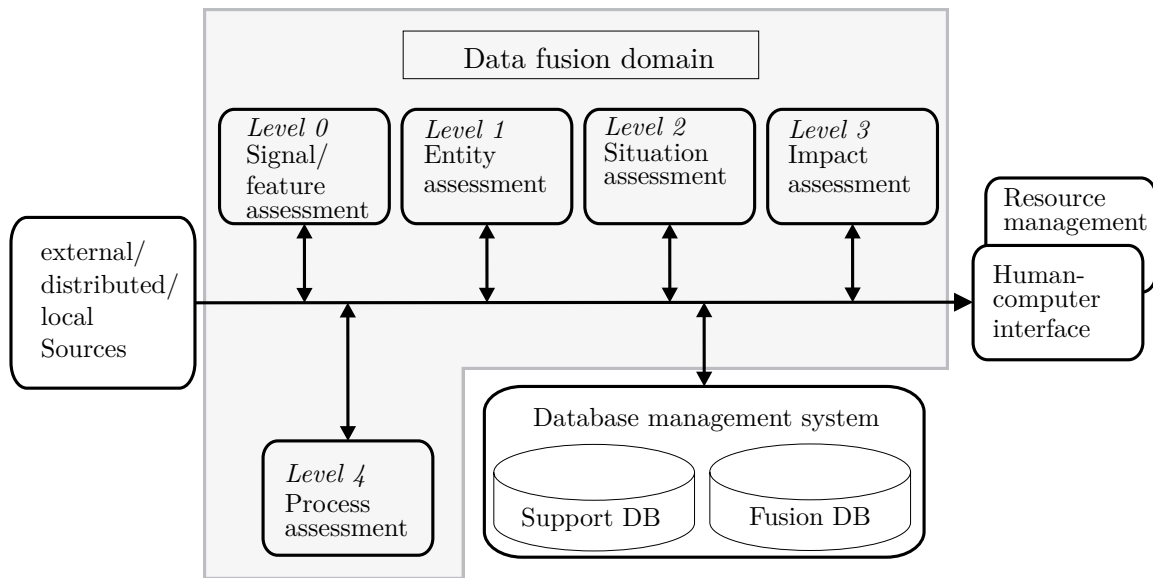


Abbildung 2.2: Prozess-Modell zur Datenfusion der Joint Directors of Laboratories (nach [53])

2.1.2 Übersicht der Prozess-Modelle

Im Zuge der fortschreitenden Entwicklung im Bereich der Datenfusion wuchs das Bedürfnis nach einer einheitlichen Konvention der Begrifflichkeiten. Folglich gründete sich die *Joint Directors of Laboratories (JDL) Data Fusion Working Group*, um das gleichnamige Prozess-Modell und ein Lexikon zur Datenfusion zu entwickeln [62].

Das Modell besteht aus einer zweischichtigen Hierarchie und beschreibt die Prozesse, Funktionen und Techniken in der Datenfusion. In der in [53] empfohlenen Modellrevision von [14] wird auf der oberen Ebene der generelle Prozessablauf der Datenfusion beschrieben (siehe Abbildung 2.2). Dieser beginnt beim Dateneinzug durch die Sensor- oder Datenquellen, führt über die Vorverarbeitung und den fünf Unterprozessen zur Datenprozessierung sowie dem Datenmanagement und der abschließenden Mensch-Maschinen-Schnittstelle. Die Mensch-Maschinen-Schnittstelle ermöglicht die Beeinflussung des Fusionsprozesses durch den Benutzer und die Präsentation der Resultate.

In der zweiten Hierarchieebene werden die spezifischen Funktionen und Methoden der fünf Unterprozesse im Detail aufgeführt und beschrieben:

Level 0: *Signal/feature assessment* Bestimmung oder Schätzung von Signalen und Features oder deren Zustände. Signale und Features stellen hierbei ein Muster dar, welches in Abhängigkeit zur Messung oder Beobachtung steht.

Level 1: *Entity assessment* Bestimmung oder Schätzung von Objektzuständen, wobei ein Objekt beispielsweise als eine eigenständige Einheit definiert werden kann.

Level 2: *Situation assessment* Bestimmung oder Schätzung von bestehenden realen Strukturen, z.B. Beziehungen und Wechselwirkungen zwischen Objekten.

Level 3: *Impact assessment* Bestimmung oder Schätzung des Kosten-Nutzen-Verhältnisses der bestimmten, geschätzten und vorhergesagten Signal-, Objekt- und Situationszustände für alle alternativen Handlungsmöglichkeiten des Beobachters.

Level 4: *Process assessment* Eigeneinschätzung der System- und Prozessperformanz und anschließende Durchführung geeigneter Maßnahmen zur Optimierung.

Die Reihenfolge der Prozessierungs-Level stellt hierbei eine Empfehlung dar. Eine einzelne Ebene kann daher auch eigenständig mit einer eigenen Datenquelle prozessiert werden. Im Folgenden werden die im JDL-Modell eingeführten Definitionen für die Begrifflichkeiten Feature/Merkmal, Objekt und Situation übernommen.

Neben dem JDL-Modell existieren noch einige Varianten oder Erweiterungen sowie weitere eigenständige Modelle zur Datenfusion. Das *Input/Output Model* von Dasarathy kann als eine Verfeinerung des JDL-Modells gesehen werden, indem es die Fusionsfunktionen nach dem Typ der Eingangs- und Ausgangsdaten kategorisiert (z.B. Signale, Features, Objekte) [26]. Hierdurch können beispielsweise die JDL-Ebenen weiter unterteilt werden.

Das *Omnibus* Modell stellt eine Kombination des Modells von Dasarathy und der JDL dar und überführt sie in ein *OODA* Modell (*observe, orient, decide, act*) [8].

2.2 Sensorsysteme

Die Notwendigkeit von sensorgestützten Assistenzsystemen in der Fahrzeugführung wurde bereits im Kapitel 1 verdeutlicht. Bereits bestehende Systeme oder Verfahren werden im Folgenden aufgeführt und gegenübergestellt.

2.2.1 Erfassung mit homogenen Sensoren

Die Verwendung eines einzelnen Sensors bzw. mehrerer homogener Sensoren ohne einen entsprechenden Überlappungsraum stellt die einfachste Form der sensorgestützten Fahrzeugumfelderfassung dar. In Serienprodukten finden derzeit größtenteils Radarsensoren sowie Mono- und IR-Kameras ihren Einsatz. An zukünftigen Sensorsystemen wie Laserscanner [3, 68, 32, 46] oder Stereokameras [21, 58, 80] wird mit Hochdruck geforscht bzw. stehen diese kurz vor dem Serieneinsatz.

Der Vorteil von homogenen Sensorsystemen besteht in einem einheitlichen Datenformat sowie, z.B. im Falle von baugleichen Sensoren, in identischen Messfrequenzen. Hierdurch wird eine Fusion der Sensordaten auf einer unteren Verarbeitungsebene vereinfacht.

Bei der Verwendung eines einzelnen Sensors ergeben sich Nachteile durch die Verwundbarkeit des Sensorsystems bei einem möglichen Sensorausfall und durch eine Beschränkung des Sichtbereiches. Als Konsequenz können mehrere Sensoren mit einem gemeinsamen Überlappungsbereich verwendet werden. So werden z.B. häufig ein Nah- und ein Fernbereichsradar in und entgegengesetzt der Fahrtrichtung eingesetzt. Die hohe Abhängigkeit zum Sensorverfahren und damit eine Anfälligkeit für verfahrensbedingte Messstörungen bleibt jedoch bestehen.

2.2.2 Erfassung mit heterogenen Sensoren

Durch den Einsatz eines heterogenen Sensorsystems sollen die sensorspezifischen Nachteile im optimalen Fall kompensiert sowie die Stärken der verschiedenen Sensortypen kombiniert werden. So wird auch im aktuellen EU-Projekt zur Unfallvermeidung durch intelligente Fahrzeuge, *interactIVe* (2010 - 2013), der Datenfusion heterogener Sensoren die zentrale Rolle in zukünftigen sicherheitsrelevanten Verkehrsentwicklungen zugeschrieben [2].

Aufgrund der Heterogenität der Sensordaten ist eine Rohdatenfusion zumeist nicht mehr möglich. Bei einem häufig verwendeten Fusionsprinzip gehen die Sensoren mit unterschiedlichen Aufgaben und Gewichtungen in die Fusion ein. Ein Hauptsensor bestimmt dabei die Objekthypothesen und die weiteren Sensoren dienen innerhalb einer kooperativen Fusion zur Verifikation der Hypothesen bzw. sind für ergänzende Aufgaben wie Objekterkennung und Klassifikation zuständig. Entsprechende Sensorpaarungen, die bereits eingesetzt oder untersucht werden, bestehen beispielsweise aus einem Laserscanner und einer Kamera [113, 86, 79], einem Laserscanner und einer Stereokamera [84, 67] sowie Radar und Stereo [31] und Laser und Radar [13].

Durch die Sensorredundanz können Falschdetektionen besser erkannt werden. Aufgrund der Ungleichgewichtung der Sensoren oder der mit der Sensorkooperation verbundenen Bestimmung unterschiedlicher Messgrößen können jedoch die vom Hauptsensor verursachten Fehldetektionen nicht mehr erkannt werden.

Daher wird in dieser Arbeit die Verwendung eines konkurrierenden Sensorsystems vorgezogen. Die Sensoren werden als gleichgestellte Hypothesengeber betrachtet, um durch die Auflösung von widersprüchlichen Sensoraussagen sowohl Falsch- als auch Fehldetektionen besser zu erkennen. Beispiele für diese Sensorkonkurrenz in der Fahrzeugumfelderfassung stellen das aus einem Laserscanner und einer Kamera bestehende System in [18] dar. Ein konkurrierender Sensoraufbau aus einem Laserscanner und einem Stereokamerasystem lassen sich z.B. in [15] und [66] finden.

Als Referenzsensor wurde für diese Arbeit ebenfalls ein konkurrierendes System bestehend aus einem Mehrzeilen-Laserscanner und einem Stereokamerasystem gewählt. Durch die gegenseitige Ausgleichung ihrer Schwächen (siehe Tabelle 2.1) und der Ähnlichkeit ihrer Rohdaten stellen die beiden Sensorverfahren eine ideale Kombination dar.

Laut einer Studie [48] soll zudem der aktuell hohe Marktpreis für Laserscanner im Fahrsicherheitsbereich in den kommenden Jahren drastisch sinken. Mit dem bevorstehenden Serieneinsatz von leistungsfähigen Stereokamerasystemen, z.B. bei Daimler [37], dürften sich auch hier die Kosten reduzieren.

Merkmal	Stereokamera	Mehrebenen-Laserscanner
Messbereich	großer Sichtbereich, horizontal ca. 40° und vertikal ca. 30°	horizontal max 110° , Schwäche in Höhenauflösung mit max. 3.2°
Auflösung	hoch, > 1 MPixel	gering, < 5000 Punkte
Signalquelle	passiver Sensor (abhängig von Beleuchtung)	aktiver Sensor (unabhängig von Beleuchtung)
Genauigkeit	mit Entfernung quadratisch abnehmende Distanzgenauigkeit	nahezu distanzunabhängige Messgenauigkeit
Reichweite	ca. 4 - 80 m	ca. 1 - 150 m
Messgrößen	räumliche Information, bildhafte Information	räumliche Information, eingeschränkt bildhafte Information
Messfrequenz	stark abhängig vom Matching-Verfahren und der Bildgröße, ca. 20 Hz	max. 50 Hz

Tabelle 2.1: Gegenüberstellung der Vor- und Nachteile der Messverfahren von Stereokamerasystem und Mehrebenen-Laserscanner in der verwendeten Konstellation

2.3 Konzepte zur Sensorfusion

Der erste systematische Versuch auf den Automotive-Bereich zugeschnittene Konzepte der Sensordatenfusion zu entwickeln, erfolgte im EU-Projekt PREVENT ProFusion1 und 2 (Laufzeit 2004 - 2008) [83]. Die hierin zusammengefassten und entwickelten Konzepte stellen auch in den darauf folgenden aktuellen Großprojekten wie *interactIVe* (2010 - 2013) [2], INTERSAFE-2 (2008 - 2011) [42] und HAVEit (2008 - 2011) [56], die sich im Themenfeld der sensorgestützten Erhöhung der Verkehrssicherheit bewegen, den Stand der Technik dar.

Vier wesentliche Fusionskonzepte wurden hierbei aufgeführt und untersucht ([83] S. 25 ff):

- *Early Fusion*
- *Track-level Fusion*
- *Multi-level Fusion and Fusion Feedback*
- *Grid based Fusion*

In der *Early Fusion* werden die Daten von mehreren auch heterogenen Sensoren auf einer möglichst frühen Verarbeitungsstufe fusioniert. Als Eingabe gehen hierbei Sensordaten in Form von Rohdaten (Signalen) oder Features ein (siehe JDL-Modell in Abschnitt 2.1.2). Die Sensordaten aufeinander folgender Messungen besitzen keine Zuordnung zur Bestimmung von Bewegungsverläufen (*Tracking*). Durch die einheitliche Verarbeitung der frühzeitig fusionierten Daten wird eine Steigerung der Zuverlässigkeit, Robustheit und Konsistenz der Umfelderfassung erwartet.

Die Fusion von zeitlich aggregierten Sensordaten auf der höheren Track-Ebene wird gemeinhin als *Track-level Fusion* bezeichnet.

Die Kombination der *Early Fusion* und der *Track-level Fusion* fokussiert die *Multi-level Fusion*. Der Ansatz des *Fusion Feedback* stellt eine Untermenge der *Multi-level Fusion* dar. Hierbei werden die Ergebnisse der *Track-level Fusion* an die Datenverarbeitung der einzelnen Sensoren zurückgeführt, um beispielsweise die Objekterkennung im nächsten Messschritt zu unterstützen.

Als *Grid based Fusion* wird die diskretisierende Zusammenfassung von räumlichen Informationen mehrerer Sensoren in einem zumeist zwei-dimensionalen planaren Raster zur Umweltdarstellung verstanden. Aus der so bestimmten fusionierten Belegungs- oder Freiraumkarte werden wiederum Objekt- oder Track-Informationen abgeleitet.

2.4 Multilevel-Multisensor-Fusion

Das in dieser Arbeit vorgestellte Multilevel-Multisensor-Fusion Framework – MMFF zur Fusion von Sensordaten in der Fahrzeugumfelderfassung ist in der Lage, die Sensorinformation auf jeweils einer der drei verschiedenen Prozessierungsebenen, im Folgenden als *Low*, *Mid* und *High Level* bezeichnet, zusammenzufassen. Durch die gleichzeitige Fusion identischer Sensordaten auf den verschiedenen Ebenen können die Auswirkungen auf die Genauigkeit und Zuverlässigkeit der fusionierten Daten in Form der jeweils resultierenden Objekt-Tracks untersucht und verglichen werden. Im Allgemeinen wird hierbei durch die Reduktion der Daten in der fortschreitenden Prozessierung mit einer gleichzeitigen Informationsabnahme gerechnet [63]. Demnach müsste

auch die Fusion von Sensordaten einer höheren Abstraktionsebene, aufgrund der reduzierten Informationsdichte, in einer vergleichsweise ungenaueren Objekt- und Track-Beschreibung resultieren. Andererseits nimmt mit steigender Abstraktion der Sensordaten die Datengröße und dementsprechend die nötige Übertragungsbandbreite und Rechenleistung zur Prozessierung ab.

Des Weiteren wird vermutet, dass im Zuge der Datenverarbeitung Fehler induziert werden, die sich als gegensätzliche Sensoraussagen innerhalb der Fusion darstellen und die im Zuge der Sensorkonkurrenz aufgelöst werden können. Hierbei sollten sich auf einer höheren Abstraktionsebene die Verarbeitungsfehler aller vorangegangenen Prozessschritte bestimmen und beseitigen lassen. Demnach müsste durch eine Kombination der Fusionsebenen eine optimale Zustandsbestimmung der Objekte durch eine Fusion auf unterster Ebene und eine Beseitigung der Prozessierungsfehler durch eine Fusion auf der oberen Ebenen zu erzielen sein.

Abbildung 2.3 stellt die in dieser Arbeit erfolgte MMFF-Realisierung unter Einbeziehung eines Laserscanners und einer Stereokamera dar. Die Definition der drei Fusionsebenen orientiert sich hierbei an der üblichen Kategorisierung (siehe beispielsweise [53] S. 7 ff, bzw. [63]), wonach auf der untersten Ebene (Low Level) die direkten Sensordaten (*JDL-signals*), hier räumliche Punktdaten, fusioniert werden. Die Sensordaten stellen dabei keine Rohdaten dar, da sie in der Vorverarbeitung (Pre Level) bereits kalibriert und in ein gemeinsames räumliches Koordinatensystem transformiert wurden. In Anlehnung an die Kategorisierung der Fusionskonzepte in Abschnitt 2.3 würde die Low-Level-Fusion dem Konzept der *Early Fusion* entsprechen.

Auf der mittleren Ebene (Mid Level) werden die Punktdaten in Form von Segmenten (*JDL-features*) geclustert. Zusammengehörige Segmentgruppierungen, die jeweils ein gemeinsames reales Objekt abbilden, werden anschließend zu Objekthypothesen zusammengefasst (*Merging* und *Shaping*) und können wiederum fusioniert werden. Im Folgenden werden Objekthypothesen auch kurz als Objekte bezeichnet.

Die Objekte der zeitlich aufeinander folgenden Messschritte werden auf der obersten Ebene (High Level) über das Tracking einander zugeordnet. Aus der Objektverfolgung resultieren die Objekt-Tracks (kurz Tracks). Durch die Integration eines Bewegungsmodells im Tracking lassen sich zudem Vorhersagen über die zukünftige Objektbewegung treffen, die als Grundlage zur Situationsbeschreibung und zur Ableitung von Handlungsentscheidungen (*JDL-decisions*) dienen. Die Tracks bzw. deren aktuelle Objektzustände der verschiedenen Sensoren werden auf der obersten Ebene fusioniert. Somit fällt die High-Level-Fusion in die im Abschnitt 2.3 vorgestellte Kategorie der *Track-level Fusion*.

Wird das MMFF in einer zentralen Fusionsarchitektur betrieben, so werden die Sensordaten nur einmal pro Messschritt auf der entsprechenden Ebene fusioniert. Bei einem hierarchischen Aufbau des MMFF können die Sensordaten hingegen auf mehreren Ebenen fusioniert werden.

Bezogen auf das JDL-Prozess Modell (siehe Abschnitt 2.1.2) sind die hier definierten Ebenen der Vorverarbeitung und die untere Verarbeitungsebene dem Level 0 (*Signal/feature assessment*) und die mittlere und obere Ebene dem Level 1 (*Entity assessment*) zuzuordnen. Die Level 2 - 4 wurden nicht berücksichtigt, da sie die zu untersuchende Fragestellung nicht betreffen.

Das MMFF wurde innerhalb des pluginbasierten Bildverarbeitungs-Frameworks *MOOVIE*¹ entwickelt, um einen modularen Aufbau sowie die Steuerung der Prozessierung und die Visuali-

¹MOOVIE ist ein an der HU Berlin von Björn Pilz und Jürgen Wohlfeil entwickeltes SW-Framework zur Bildverarbeitung

sierung der Daten zu ermöglichen.

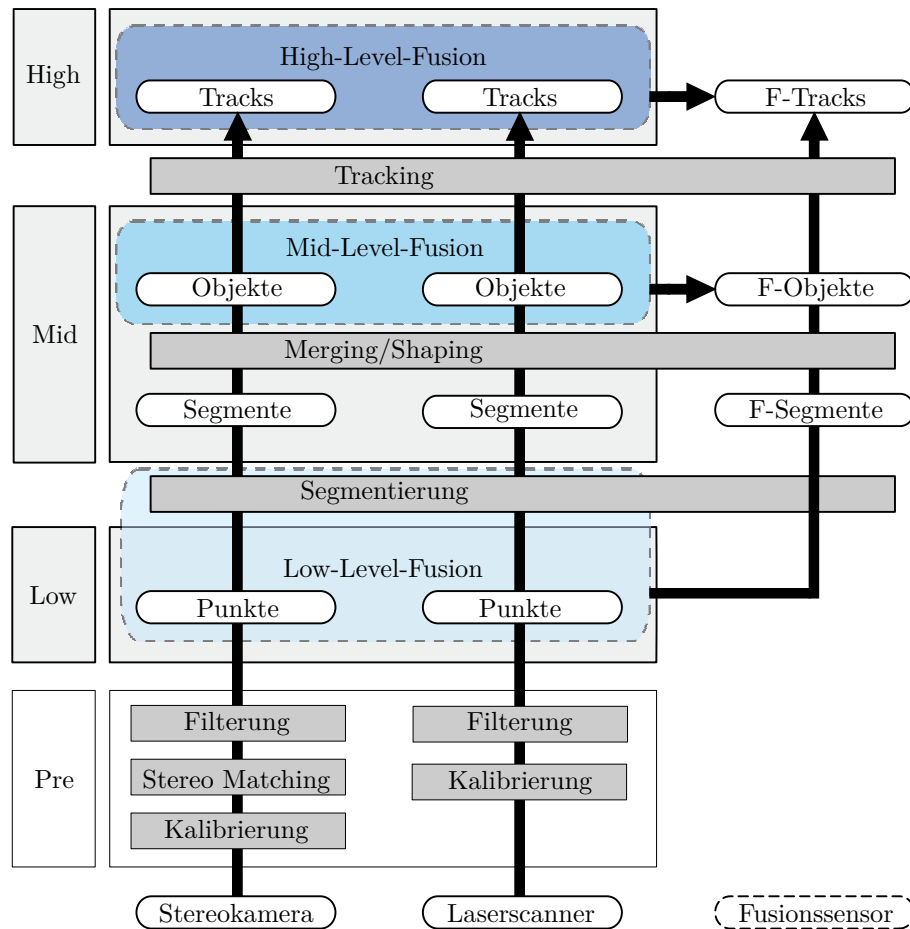


Abbildung 2.3: Framework zur Multilevel-Multisensor-Fusion für das vorgestellte Sensorsystem bestehend aus Stereokamera und Laserscanner

Um neben dem Stereokamerasystem und dem Laserscanner noch weitere Sensoren und Sensortypen in das Framework einbinden zu können, wurde eine allgemeinere Beschreibung gewählt (siehe Abbildung 2.4). Aufgrund der Multilevel-Architektur können ebenfalls Sensoren integriert werden, die nur Daten auf höheren Ebenen liefern. So können beispielsweise die Punktdaten von Laserscanner und Stereokamera auf der unteren Ebene fusioniert werden und die Objektdaten eines Radarsystems gehen entsprechend auf der mittleren oder oberen Ebene in die Fusionsdaten mit ein (hierarchische Fusion).

Die Eigenschaften des Multilevel-Multisensor-Fusion Frameworks lassen sich somit wie folgt zusammenfassen:

skalierbar Die Anzahl der einzubindenden Sensoren ist nicht beschränkt.

typoffen Über eine entsprechende Abstraktion der Prozessierungs- und Fusionsverfahren gibt

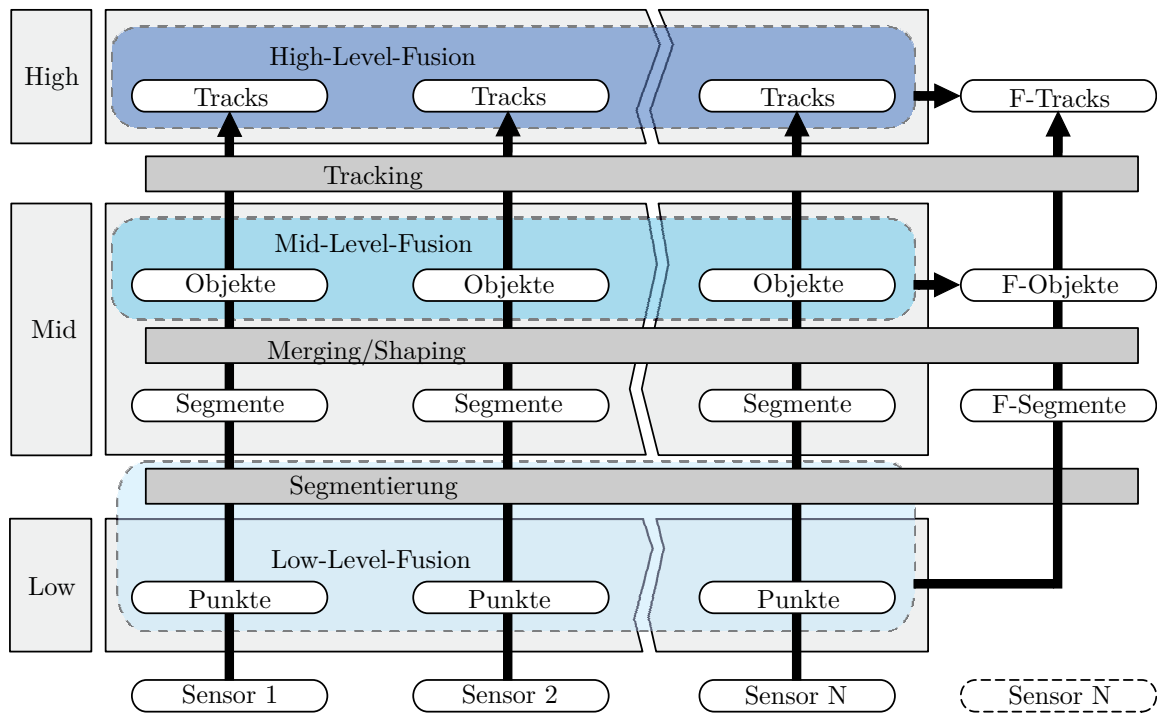


Abbildung 2.4: Allgemeines Multilevel-Multisensor-Fusion Framework für N Sensoren

es keine Beschränkung auf festgelegte Sensorverfahren.

hierarchisch Die Prozessierung der Sensordaten kann dezentral erfolgen. Für mehr als zwei eingehende Sensoren können die Fusionsverfahren der mittleren und oberen Ebene ebenfalls dezentral an mehreren Fusionsknoten ausgeführt werden.

konkurrierend Die unabhängigen Sensoren tragen durch Informationsredundanz zur Verminderung der Unsicherheit und Steigerung der Detektionsgenauigkeit und Zuverlässigkeit bei.

komplementär In den ausschließlichen Sichtbereichen eines Sensors werden die Daten ohne eine Sensorfusion bis zur obersten Ebene weiterverarbeitet.

Um die verschiedenen Vorteile der drei Fusionsebenen zu kombinieren und deren Nachteile zu kompensieren ist das MMFF durch eine sequentielle Kombination der Fusionsebenen innerhalb eines Messschrittes erweitert worden (siehe hierzu Abschnitt 6.4).

Das MMFF kann somit als eine Erweiterung des in [83] und [96] formulierten Konzeptes der *Multi-level Fusion* angesehen werden. Die in [96] vorgestellte *Multi-level Fusion* versucht in Abhängigkeit der eingehenden Sensoren und der Anwendungsspezifik die Genauigkeit der Objektbestimmung zu erhöhen, indem die Datenfusion auf jeweils einer bestimmten optimalen Fusionsebene durchgeführt wird. Eine parallele Fusion zur Zusammenfassung der jeweiligen Vorzüge der einzelnen Ebenenfusionen findet nicht statt. Der Begriff Multi-Level bezeichnet dort das Zusammenführen von Daten mehrerer Sensoren aus unterschiedlichen Verarbeitungsstufen nicht

die parallele Fusion auf verschiedenen Ebenen wie beim MMFF. Ein weiterer Unterschied besteht in der gegenseitigen Unabhängigkeit der eingebundenen Sensoren und ihrer Prozessierung im MMFF, um eine gleichberechtigte Konkurrenz zu gewährleisten. Hierdurch ist das MMFF durch Fusionsverfahren auf den höheren Verarbeitungsebenen in der Lage, Fehl- und Falschinformationen gezielt zu detektieren und zu reduzieren.

Der Ansatz des *Fusion Feedbacks* forciert eine Rückkopplung der Fusionsergebnisse (Tracks) von der Track-Ebene auf die darunter liegenden Verarbeitungsebenen einiger vermeintlich schwächer detektierender Sensoren. Durch die fusionierte Track-Information soll die Prozessierung rückwirkend oder im nächsten Zeitschritt hinsichtlich Fehl- und Falsch-Detektionen überprüft und korrigiert werden. Bei einer Rückkopplung in den nächsten Zeitschritt besteht hierbei die Gefahr, dass sich Fehler auf der Track-Ebene durch die Abbildung auf eine untere Verarbeitungsebene wieder selbst injizieren können. Im Vergleich zum MMFF besteht zudem keine parallele Fusion auf unterster Ebene zur Optimierung der Objektbeschreibung.

Die *Grid based Fusion* verliert im Vergleich zur MMFF durch die diskretisierende Zellendarstellung an Informationen und Beschreibungsgenauigkeit. Zudem wird entweder auf Feature- oder Track-Ebene fusioniert, wodurch sich Nachteile in der Genauigkeit der Objektbestimmung oder der Fehlerreduzierung ergeben können.

2.5 Zusammenfassung

Ziel des Kapitels war es, eine Einführung in die Begrifflichkeiten und Methoden der sensorgestützten Fahrzeugumfelderfassung und der damit eng verbundenen Sensordatenfusion zu geben. Des Weiteren wurde anhand der technologischen Evolution, die sich vom Einzelsensor bis hin zum heterogenen Fusionssystem entwickelt, die aktuell bestehenden Sensorsysteme bzw. das in der Arbeit entwickelte Stereokamera-Laserscanner-System eingeordnet. Es erfolgte zudem die Vorstellung des Multilevel-Multisensor-Fusion Frameworks, dessen Einordnung in das gebräuchliche JDL-Prozess-Modell und ein Vergleich des postulierten Innovationsschubs zum Stand der Technik.

3 Sensorsystem

Zur praktischen Erprobung und Validierung des vorgestellten MMFF wurde ein Referenzsensor bestehend aus einem Mehrzeilen-Laserscanner und einem Stereokamerasystem zur Umfelderkennung am Fahrzeug entwickelt. Im Folgenden wird der Aufbau des Sensorsystems, die Funktionsweise und Kalibrierung der Sensoren sowie die Vorverarbeitung der Sensordaten beschrieben.

3.1 Aufbau des Sensorsystems

Das dieser Arbeit zugrunde liegende konkurrierende Sensorsystem besteht aus einem 4-Zeilen-Laserscanner und beidseitig davon angebrachten Stereokameras. Die Sensoren werden auf einem Profil im Vorderbereich eines Messfahrzeugs montiert, wodurch ihr Erfassungsbereich in Fahrtrichtung zeigt (siehe Abbildung 3.1(b)). Die Profilhöhe über der Fahrbahn beträgt hierbei ca. 80 cm. Die Daten gelangen über die Ethernet-Schnittstelle zur Weiterverarbeitung und Aufzeichnung in einen fahrzeugintegrierten Rechner. Zur zeitlichen Synchronisierung der angebundenen Sensoren dient eine im Rechner integrierte GPS-Funkuhr als NTP-Zeitserver. Das übergeordnete Koordinatensystem des Sensorsystems (*SKS*) entspricht gleichzeitig dem Koordinatensystem des Stereosystems (*SKKS*) bzw. dem Kamerakoordinatensystem der linken Kamera (siehe Abbildung 3.3).

3.1.1 Spezifikation des Stereosystems

Aufbau

Das Stereosystem wurde im Rahmen der Arbeit entwickelt und setzt sich aus zwei Kameras des Typs *PicSight P141M Smart* von *Leutron Vision* zusammen. Die Kameras besitzen pro Pixel einen Dynamikumfang von 8-Bit bei einer maximalen Auflösung von 1392×1040 Pixeln. Zudem verfügen die Kameras über eine integrierte Recheneinheit, die zur Anbindung an den NTP-Server genutzt wird, um die Bilder direkt bei der Aufnahme mit einem präzisen Zeitstempel zu versehen.

Bei einer sich schnell veränderten Umgebung, ist eine möglichst zeitgleiche Bildaufnahme beider Stereokameras für eine korrekte Berechnung der Tiefeninformationen notwendig. Die Eigenbewegung der Kameras bzw. die Veränderung der Szenerie zwischen den zeitversetzten Einzelaufnahmen eines Stereobildpaares können die Parallaxeninformation der Bildpunktkorrespondenzen stark verzerren. Um den Effekt zu reduzieren, wird die Bildaufnahme der beiden Kameras über einen externen elektrischen Impuls möglichst zeitgleich getriggert. Versuche mit einer optischen digitalen Uhr haben dabei ergeben, dass der zeitliche Aufnahmeversatz unter $10 \mu s$ liegt.

Aufgrund einer gewünschten Erfassungstiefe von 5 m - 60 m wurde in Abhängigkeit der maximalen Disparitäten-Suchlänge von 128 Pixeln, einer Subpixeltiefe von 16 und der verwendeten

3 Sensorsystem

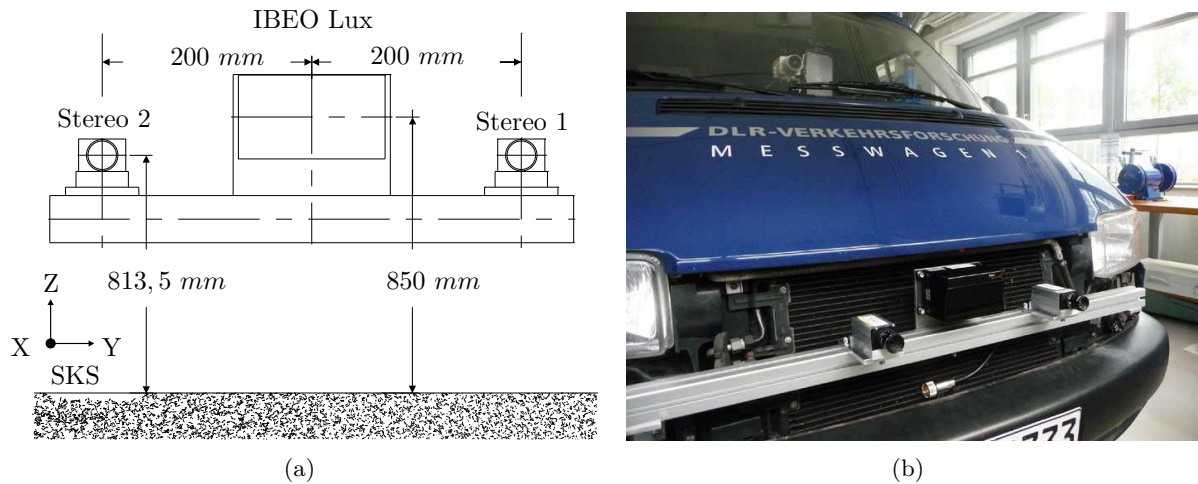


Abbildung 3.1: Schematische Darstellung des Sensoraufbaus (a); Befestigung des Sensoraufbaus am Messfahrzeug UTRaCarTM(b)

Kamerakonstante c von 12.5 mm eine Basis B zwischen den Projektionszentren der Kameras mit einer Länge von ca. 40 cm gewählt. Abbildung 3.2 veranschaulicht hierzu als Histogramm-Darstellung die jeweiligen Erfassungsbereiche für unterschiedliche Basislängen mit der gewählten Stereokonfiguration, wobei sich im Bereich von $B = 40$ cm der gewünschte Erfassungsbereich ergibt. Als Annahme wird von einer Gleichverteilung der Einzelmessungen über dem Disparitätensuchraum von $128 \times 16 = 2048$ möglichen Disparitätsstufen ausgegangen.

Eine Reduktion der Laserscannerdaten ist aufgrund der hohen Messgenauigkeit und der geringen Punktdichte nicht nötig.

Funktionsweise

Für die Stereobildverarbeitung wird in dieser Arbeit vom Stereo-Normalfall ausgegangen. Hierbei sind die beiden Projektionszentren O' und O'' der Kameras über die Basis B verbunden. Die Bilder bzw. die entsprechenden Kamerakoordinatensysteme der beiden Kameras werden durch die *Rektifizierung* dahingehend transformiert, dass ihre Aufnahmerichtungen orthogonal zur Basis und parallel zueinander liegen. Zusätzlich sollen die *Epipolarlinien* der Stereobilder parallel zur Basis verlaufen. Die jeweiligen Epipolarlinien ergeben sich hierbei bei einem Schnitt der zugehörigen *Epipolarebene* mit den beiden Bildebenen, wobei eine Epipolarebene durch die beiden Projektionszentren O' und O'' und einen in beiden Bildern sichtbaren Objektpunkt aufgespannt wird. Des Weiteren wird für beide Kameras, wie in Abbildung 3.3 dargestellt, eine gemeinsame Kamerakonstante c , die dem Abstand vom Projektionszentrum zum Bildmittelpunkt entspricht, festgelegt (siehe beispielsweise [65] S. 28 ff).

Durch diese normalisierte Epipolargeometrie liegen die Bildpunktkorrespondenzen eines Stereobildpaares jeweils auf der gleichen Bildzeile in beiden Bildkoordinatensystemen (BKS), wodurch sich die automatisierte Korrespondenzsuche vereinfacht. Für zwei korrespondierende Bildpunkte des linken (x', y') und rechten Kamerabildes (x'', y'') strebt somit die Y-Parallaxe gegen

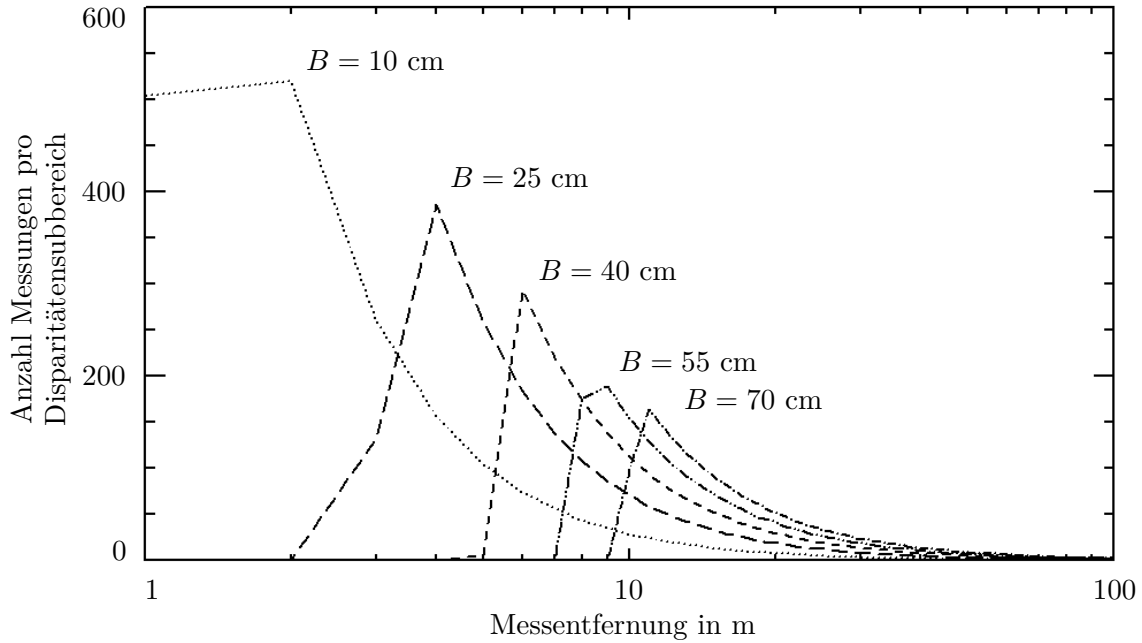


Abbildung 3.2: Erfassungsbereiche in Abhängigkeit zur Basis B bei einer maximalen Disparitäten-Suchlänge von 128 Pixeln, einer 16-fachen Subpixelauflösung und einer Kamerakonstante von 12.5 mm; die jeweiligen 128×16 Einzelmessungen liegen gleichverteilt über dem Disparitätensuchraum

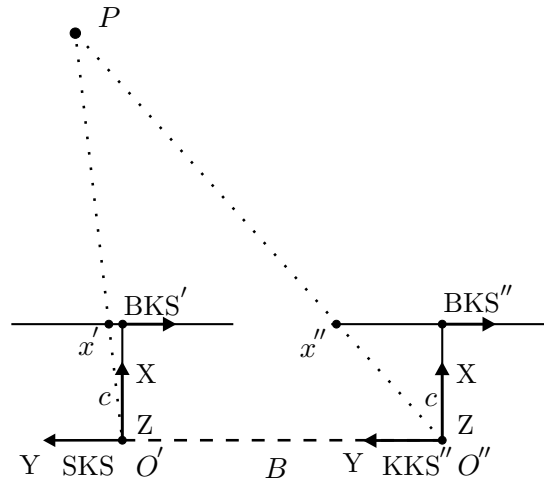


Abbildung 3.3: Stereokamerasystem im Normalfall: Objektpunkt P wird auf der X-Achse der Bildkoordinatensysteme der linken (BKS') und rechten Kamera (BKS'') auf x' und x'' abgebildet; Basis B liegt zwischen den beiden Projektionszentren O' und O'' ; Achsenrichtungen beider Kamerakoordinatensysteme (KKS) sind parallel zueinander; Kamerakonstante c ist für beide Kameras identisch

Null und die X-Parallaxe entspricht der gesuchten Disparität d :

$$d = x'' - x' . \quad (3.1)$$

Die Berechnung der Koordinaten (X, Y, Z) eines Objektpunktes P über die abbildenden Bildsensorkoordinaten x' und y' des linken Bildes vereinfacht sich dementsprechend zu:

$$\begin{aligned} X &= \frac{Bc}{d} \\ Y &= -X \frac{x'}{c} \\ Z &= X \frac{y'}{c} . \end{aligned} \quad (3.2)$$

Die Überführung der Stereobilder durch die Rektifizierung in den Stereo-Normalfall wird durch eine projektive Transformation vollzogen. Hierbei werden zunächst die 2D-Koordinaten eines Bildes über die entsprechende Kamerakonstante in den 3D-Raum überführt. Diese Bildvektoren lassen sich durch die jeweilige über die relative Orientierung bestimmte (siehe Abschnitt 3.2.1) Rotation zum anderen Bild in die Normallage drehen, um anschließend wieder über eine gemeinsame Kamerakonstante in die rektifizierten 2D-Bildkoordinaten abgebildet zu werden (Näheres siehe z.B. [71] S. 257).

Die automatisierte Bestimmung der Bildpunktkorrespondenzen innerhalb eines Stereobildpaares wird als Stereo-Matching bezeichnet. Die Matching-Verfahren werden hierbei oft, so auch in [95], in lokale und globale Verfahren unterteilt.

Bei den lokalen Verfahren wird die Korrespondenzwahrscheinlichkeit zweier zu vergleichenden Bildpunkte im Stereobildpaar (Stereobildpunkte) über die Ähnlichkeit ihrer lokalen Bildpunkt-Nachbarschaften (zumeist quadratisches Fenster) bestimmt. Die Bildpunktpaarung mit der höchsten Korrespondenzwahrscheinlichkeit bestimmt entsprechend die Höhe der Bildpunktverschiebung auf der X-Achse (Disparität).

Die globalen Verfahren verwenden hingegen noch zusätzliche Annahmen für die Korrespondenzanalyse, wie die Eineindeutigkeit der Bildpunktkorrespondenzen (*uniqueness*), die kontinuierliche Änderung der Disparitäten im überwiegenden Teil des Bildraums (*continuity* oder *smoothness*) oder eine einseitige Verdeckung im Objektraum (*occlusion*) und nutzen diese zur Optimierung einer globalen Kostenfunktion [117].

Im Rahmen der Arbeit wurden diverse Verfahren implementiert und evaluiert. Als lokaler Vertreter wurde aufgrund seiner schnellen Berechnung und akzeptablen Genauigkeit der fensterbasierte *Block Matcher* (BM) in der Implementierung von *openCV* ausgewählt. Für die Berechnung eines dichteren Disparitätenbildes wird der *Semi Global Block Matcher*, eine vereinfachte Implementierung des *Semi Global Matcher* (siehe [55]), von *openCV* eingesetzt, der jedoch eine um Faktor zehn langsamere Rechenzeit benötigt als der BM. Nach der Auslagerung der berechnungsintensiven Operationen der Stereobildverarbeitung (Rektifizierung, Matching, Punktbestimmung) auf eine GPU konnten für den BM bei 96 Pixel Suchtiefe und voller Auflösung

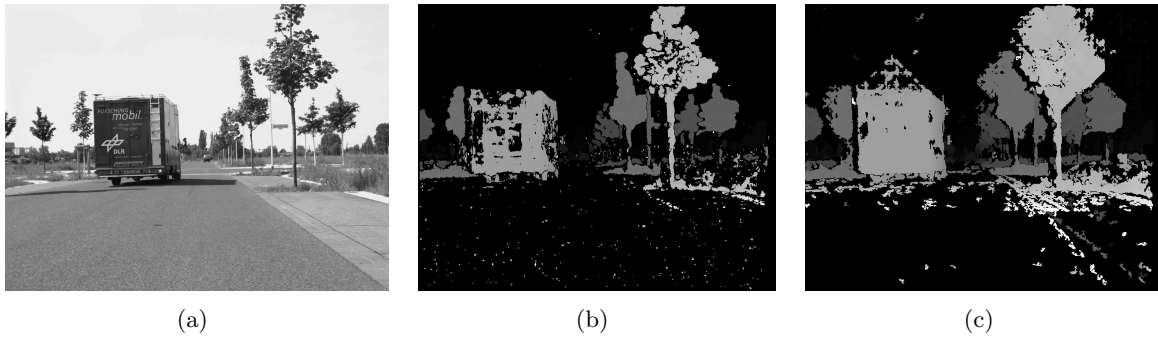


Abbildung 3.4: Stereobild der linken Kamera (a); normalisiertes Disparitätenbild durch BM-Verfahren (b); normalisiertes Disparitätenbild durch SGBM-Verfahren (c); eine abnehmende Intensität im Disparitätenbild entspricht einer ansteigenden Messentfernung; maximale Disparitätensuchlänge entspricht hierbei 96 Pixel

eine Taktfrequenz von ca. 20 Hz erreicht werden. In Abbildung 3.4 sind die Tiefenbilder beider Verfahren für eine Beispielszene dargestellt.

3.1.2 Spezifikation des Laserscanners

Aufbau

Der Mehrzeilen-Laserscanner *Lux* der Firma *IBEO* [59] arbeitet wahlweise in den Frequenzen 12.5, 25 und 50 Hz. Er verfügt neben einem Synchronisierungseingang über einen Synchronisierungsausgang, der beim Durchlaufen des Messnullpunktes (entspricht der Zielachse) einen Stromimpuls erzeugt. Dieser Impuls wird genutzt, um gleichzeitig die Stereokameras auszulösen, so dass beide Sensoren zu nahezu gleichen Zeitpunkten und im gleichen Intervall ihre Daten akquirieren. Diese von der Praxis eher abweichende ideale Synchronisierung der Sensoren wurde gewählt, um die Sensordaten ebenfalls auf der untersten Prozessierungsebene kombinieren zu können. Durch die Begrenzung der Scanfrequenz der Kameras unter Vollauflösung auf unter 20 Hz wird das Gesamtsystem daher mit 12.5 Hz betrieben.

Funktionsweise

Der Laserscanner bestimmt die Messstrecke r über die Laufzeit zwischen Signalausendung und dem Empfang der Reflexion (*Time of Flight*). Mit vier horizontal übereinander gelagerten Scan-Ebenen wird die Umgebung mit einem maximalen horizontalen Öffnungswinkel von 110° und einer vertikalen Öffnung von 3.2° erfasst. Neben der Streckenlänge werden über die Ablenkung des Laserstrahls von der Zielachse ein horizontaler Messwinkel φ sowie ein von der Messebene abhängiger vertikaler Messwinkel θ für jeden Messpunkt P bestimmt (siehe Abbildung 3.5).

Die Umrechnung dieser polaren Werte in kartesische Messkoordinaten des Laserscanner-Koordinatensystems (*LKS*) erfolgt mit:

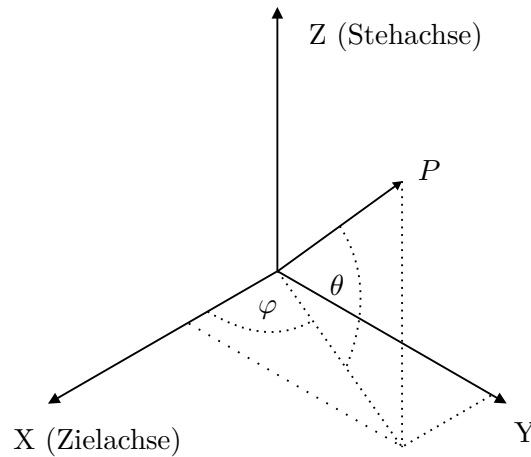


Abbildung 3.5: Vertikale θ und horizontale Ablenkung φ des Laserstrahls von der Zielachse im polaren Koordinatensystem

$$\begin{aligned} x &= r \cos \varphi \cos \theta \\ y &= r \sin \varphi \cos \theta \\ z &= r \sin \theta . \end{aligned} \tag{3.3}$$

3.2 Kalibrierung der Sensoren

Die Sensorkalibrierung umfasst die Bestimmung und Entfernung des systematischen Fehleranteils sowie die Schätzung des zufälligen Fehleranteils in der Messung. Zur Bestimmung der beiden Fehlertypen müssen diese für jeden Sensor modellhaft beschrieben werden. Die Modellierung des systematischen Fehlers basiert zumeist auf einer physikalischen Beschreibung des Messverfahrens bzw. damit einhergehender Stör- und Fehlereinflüsse. Der zufällige Fehler wird über eine passende Wahrscheinlichkeitsverteilung geschätzt. Zusätzlich umfasst die Sensorkalibrierung die Überführung der räumlichen Daten unterschiedlicher Sensoren in ein gemeinsames Koordinatensystem.

3.2.1 Kalibrierung der Stereokamera

Um eine möglichst optimale Messgenauigkeit zu erzielen, wird das Stereokamerasystem im Vorfeld unter Laborbedingungen kalibriert. Der Kalibrierungsprozess unterteilt sich einerseits in drei bekannte aufeinander folgende Einzelverfahren, die daher hier nur kurz beschrieben werden. Zur Bestimmung des zufälligen Fehlers wurde eine Erweiterung entwickelt.

Bestimmung des systematischen Fehlers und der Kameraorientierung

Innere Orientierung (IO) Im Zuge der inneren Orientierung werden für die beiden einzelnen Kameras auf Grundlage eines Abbildungs- und Verzeichnungsmodells die systematischen und zufälligen Messungenauigkeiten zur Abbildung eines räumlichen Objektpunktes in den Bildraum ermittelt. Hierbei werden die kameraspezifischen Parameter der Hauptpunktverschiebung (x_p, y_p) und der Kamerakonstanten (c) sowie die Parameter der radial symmetrischen Verzerrung (k_1, k_2, k_3), der radial asymmetrischen Verzerrung (p_1, p_2) [16] und die Parameter der Affinität und Scherung (b_1, b_2) [30] bestimmt. Zur Schätzung der Parameter dient eine Kalibrierwand mit bekannten räumlichen Referenzpunkten höherer Genauigkeit (siehe Abbildung 3.6). Über die Zuordnung zu den abbildenden Bildpunkten für mehrere Kamerastandpunkte kann eine Parameterschätzung über einen Rückwärtschnitt im Zuge eines Bündelausgleichs erfolgen ([71] S. 266 ff).

Relative Orientierung (RO) Die relative Orientierung beschreibt die gegenseitige Ausrichtung (drei Rotationsparameter ω, ϕ, κ) und Verschiebung (drei Translationsparameter T_x, T_y, T_z) zweier Kamerakoordinatensysteme. In dieser Arbeit wird die relative Orientierung der Stereokameras durch Bilddrehung ermittelt (siehe hierzu [65] S. 218 ff). Hierbei wird der Translationsvektor als Y-Koordinatenachse im Modellkoordinatensystem (MKS) definiert und nur die beiden gegenseitigen Drehungen werden über fünf Drehwinkel ($\omega_L, \phi_L, \kappa_L, \phi_R, \kappa_R$) bestimmt. Hierzu dienen die durch die IO normalisierten, korrespondierenden Bildpunkte (x', y' und x'', y'') eines Stereobildpaares, die jeweils den gleichen Referenzpunkt X_R abbilden. Die beiden korrespondierenden Vektoren V' und V'' werden in den Kamerakoordinatensystemen mit O', x', y' und O'', x'', y'' gebildet und bestimmen im MKS in der Mitte ihrer kürzesten Verbindung die entsprechende Modellkoordinate X_M von X_R dar. Durch die Minimierung der von der Basis und den korrespondierenden Bildvektoren V' und V'' aufgespannten Spatvolumen können die gesuchten Parameter in einer überbestimmten Ausgleichung geschätzt werden.

Absolute Orientierung (AO) Im letzten Schritt der absoluten Orientierung wird die modellhafte relative Orientierung in ein übergeordnetes Weltkoordinatensystem (WKS) durch die sieben Parameter einer Ähnlichkeitstransformation (Skalierung m_a , Rotation R_a , Translation T_a) in einer direkten Bestimmung überführt [57]. Über die Skalierung lässt sich so die reale Basislänge B zwischen den beiden Projektionszentren der Kameras O' und O'' bestimmen. Als abschließender Orientierungsschritt für das Stereosystem steht die Güte der AO, welche sich aus den Residuen zwischen den Referenz- und transformierten Modellkoordinaten X_W ergibt, für die Gesamtgüte des Systems.

Für das vorgestellte Sensorsystem liegt der mittlere Fehler (RMSE) nach der Kalibrierung typischerweise für die IO unter 1/30 Pixel im Bildbereich und für die AO unter 1 mm im Objektraum.

Da sich die Kalibrierung auch über mehrere Testfahrten als stabil erwiesen hat, wurde auf eine kontinuierliche Nachführung der RO während der Fahrt, wie z.B. in [25] vorgestellt, verzichtet.

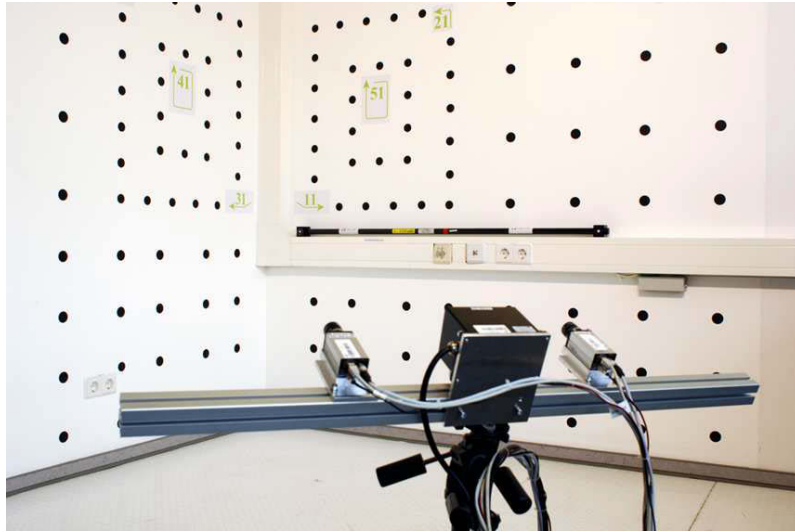


Abbildung 3.6: Kalibrierwand mit Referenzpunkten zur Bestimmung von IO, RO und AO

Bestimmung des zufälligen Fehleranteils

Die Messgenauigkeit eines Stereosystems ist zum einen von den Einzelgenauigkeiten der Stereokameras ($\sigma_{x'}$, $\sigma_{y'}$ und $\sigma_{x''}$, $\sigma_{y''}$), modelliert durch die innere Orientierung sowie von den Genauigkeiten der gemeinsamen relativen und absoluten Orientierung abhängig. Zusätzliche Fehlerquellen bei der Bestimmung der räumlichen Punkte durch das Stereoverfahren liegen in dem verwendeten Stereo-Matching (z.B. Subpixelstufen in der Disparitätsbestimmung) sowie in der Rektifizierung der Stereobildpaare. Zur besseren Abschätzung der Gesamtfehlers bzw. der Bildpunktvarianzen σ_x und σ_y wird daher die absolute Orientierung mit einbezogen, da diese als abschließender Kalibrierungsschritt die Gesamtgenauigkeit des Stereosystems repräsentiert. Um zusätzlich den Einfluss des Matching-Verfahrens auf die Messgenauigkeit mit abzubilden, werden die Modellkoordinaten \mathbf{X}_M zu den entsprechenden Referenzkoordinaten \mathbf{X}_R im Labor (siehe Abbildung 3.6) nicht wie bei der Ermittlung der absoluten Orientierung über die relative Orientierung, sondern direkt durch das Stereo-Matching und den Gleichungen 3.2 bestimmt. Die Modellkoordinaten werden anschließend über die Parameter der bereits bestimmten absoluten Orientierung mit

$$\mathbf{X}_W = \mathbf{T}_a + \mathbf{R}_a \mathbf{X}_M \quad (3.4)$$

in das Referenzkoordinatensystem als Weltkoordinaten \mathbf{X}_W überführt (Skalierung $m_a = 1$). Der Messfehler e^W im WKS kann nun als Distanz zwischen den Weltkoordinaten \mathbf{X}_W und ihren Referenzkoordinaten \mathbf{X}_R im Raum mit

$$\begin{aligned}
e_x^W &= \mathbf{X}_{R_x} - \mathbf{X}_{W_x} \\
e_y^W &= \mathbf{X}_{R_y} - \mathbf{X}_{W_y} \\
e_z^W &= \mathbf{X}_{R_z} - \mathbf{X}_{W_z}
\end{aligned} \tag{3.5}$$

bestimmt werden. Diese Fehlerdistanz wird über ihre beiden Endpunkte mit den Kollinearitätsgleichungen ([65] S. 19) aus dem Objektraum (WKS) in den Bildraum (BKS) der entsprechenden Kamera in X- Richtung mit

$$\begin{aligned}
e_x^B &= x_p - c \left(\frac{\mathbf{R}_{a11}(\mathbf{X}_{R_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a12}(\mathbf{X}_{R_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a13}(\mathbf{X}_{R_z} - \mathbf{T}_{a_z})}{\mathbf{R}_{a31}(\mathbf{X}_{R_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a32}(\mathbf{X}_{R_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a33}(\mathbf{X}_{R_z} - \mathbf{T}_{a_z})} \right. \\
&\quad \left. - \frac{\mathbf{R}_{a11}(\mathbf{X}_{W_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a12}(\mathbf{X}_{W_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a13}(\mathbf{X}_{W_z} - \mathbf{T}_{a_z})}{\mathbf{R}_{a31}(\mathbf{X}_{W_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a32}(\mathbf{X}_{W_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a33}(\mathbf{X}_{W_z} - \mathbf{T}_{a_z})} \right)
\end{aligned} \tag{3.6}$$

und in Y-Richtung mit

$$\begin{aligned}
e_y^B &= y_p - c \left(\frac{\mathbf{R}_{a21}(\mathbf{X}_{R_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a22}(\mathbf{X}_{R_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a23}(\mathbf{X}_{R_z} - \mathbf{T}_{a_z})}{\mathbf{R}_{a31}(\mathbf{X}_{R_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a32}(\mathbf{X}_{R_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a33}(\mathbf{X}_{R_z} - \mathbf{T}_{a_z})} \right. \\
&\quad \left. - \frac{\mathbf{R}_{a21}(\mathbf{X}_{W_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a22}(\mathbf{X}_{W_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a23}(\mathbf{X}_{W_z} - \mathbf{T}_{a_z})}{\mathbf{R}_{a31}(\mathbf{X}_{W_x} - \mathbf{T}_{a_x}) + \mathbf{R}_{a32}(\mathbf{X}_{W_y} - \mathbf{T}_{a_y}) + \mathbf{R}_{a33}(\mathbf{X}_{W_z} - \mathbf{T}_{a_z})} \right)
\end{aligned} \tag{3.7}$$

abgebildet. Die Varianz der Punktmessung der Stereokamera kann mit einer signifikanten Anzahl von N Fehlerdistanzen e_x^B und e_y^B als σ_x und σ_y im Bildraum bestimmt werden:

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (e_x^B)^2 \tag{3.8}$$

$$\sigma_y^2 = \frac{1}{N-1} \sum_{i=1}^N (e_y^B)^2. \tag{3.9}$$

Die Werte für σ_x und σ_y liegen bei dem verwendeten Aufbau und dem eingesetzten Block-Matcher für beide Kameras bei ca. 1/3 Pixel.

3.2.2 Kalibrierung des Laserscanners

Die Kalibrierung des Laserscanners umfasst zum einen die Bestimmung des systematischen und zufälligen Messfehlers sowie die Bestimmung der äußeren Orientierung zum Bezugskordinatensystem, in diesem Falle das SKKS der Stereokamera.

Fehlermodell

Im Folgenden soll ein geeignetes parametrisierbares Modell der systematisch und zufällig auftretenden Instrumentenfehler eines Mehrzeilen-Laserscanners aufgestellt werden. Die Ursache der systematischen Fehler liegt hierbei zumeist in der Abweichung des Geräteaufbaus von der idealgeometrischen Form sowie in der Lageveränderung der Bauteile mit der Zeit.

Der Streckenmessfehler e_r als Differenz der gemessenen Ist-Strecke r_M und der wahren Soll-Strecke r wird durch einen konstanten Anteil b , einen linearen Anteil a und einen zufälligen Anteil u , mit

$$r = ar_M + u + b \quad (3.10)$$

beschrieben.

Da der Laserscanner nicht wie bei einem klassischen polaren Messinstrument (z.B. Tachymeter) um zwei orthogonale Achsen, sondern nur um die vertikale Stehachse rotiert, wird der Kipp- und Zielachsenfehler (Definition z.B. siehe [90] S. 72 ff) dementsprechend nicht berücksichtigt. Die Achsenexzentrizität zwischen der Steh- und Zielachse sowie die Abweichung der Stehachse aus der Lotrechten werden nicht explizit als Instrumentenfehler betrachtet, sondern durch die äußere Orientierung abgebildet. Der Höhenindexfehler, der eine lineare Verzerrung der Messwerte durch einen konstanten Offset auf der vertikalen Winkelmessung verursacht, kann ebenfalls durch die äußere Orientierung abgebildet werden.

Zur Kompensation einer Abweichung vom idealen Auftreffwinkel der Laserstrahlen auf den Ablenkspiegel wird ein Skalierungsfaktor s für den vertikalen Messwinkel θ_M sowie eine additive Zufallsgröße v eingeführt:

$$\theta = s\theta_M + v. \quad (3.11)$$

Dabei hat sich gezeigt, dass die Verwendung einer gemeinsamen Skalierung einer separaten Skalierung der vier Scan-Ebenen, aufgrund eines kleineren Gesamtfehlers in der Ausgleichung, vorzuziehen ist.

Der horizontale Messwinkel φ_M wird als mittelwertfrei angenommen. Ein fehlerhaft auftretender Winkeloffset wird innerhalb der Ausgleichsrechnung durch die äußere Orientierung sowie das Messrauschen durch w abgebildet:

$$\varphi = \varphi_M + w. \quad (3.12)$$

Für die Zufallsgrößen u , v und w wird die Normalverteilung angenommen:

$$\begin{aligned} p(u) &\sim N(0, \sigma_r) \\ p(v) &\sim N(0, \sigma_\theta) \\ p(w) &\sim N(0, \sigma_\varphi) . \end{aligned} \tag{3.13}$$

Eine Betrachtung von äußeren Einflüssen auf die Messgenauigkeit wie die Beschaffenheit, Textur und Farbe des zu messenden Objektes, atmosphärische Einflüsse oder dem Auftreffwinkel des Laserstrahls, wie zum Beispiel in [47] beschrieben, wird hier nicht durchgeführt, da sie sich nur schwer, im Falle einer systematischen Natur, über ein Modell abbilden lassen, um die Messdaten entsprechend zu korrigieren.

Bestimmung des systematischen Fehlers und der Koregistrierung

Zur Kalibrierung des Laserscanners ist ein Vergleich zwischen einem Messpunkt und einer um mindestens eine Größenordnung genauer vermessenen Referenz nötig. Das Messprinzip und der Aufbau des Mehrebenen-Laserscanners erweisen sich jedoch als problematisch bei der Zuordnung einer entsprechenden Referenz. Klassische Ansätze, wie z.B. die Verwendung von optischen Messmarken oder Prüfkörpern (zumeist Kugeln) zur Bestimmung der Messgenauigkeit von terrestrischen Laserscannern (TLS) [61] oder eine optische Zuordnung der Messpunkte zu einem Referenzkamerasystem [93, 94, 110, 89], scheitern hier an der zu geringen Messwinkelauflösung, der geringen Anzahl der Messebenen sowie der Messung oberhalb des sichtbaren Wellenlängenbereiches.

Im Folgenden wird ein neu entwickelter Ansatz vorgestellt, der innerhalb eines Rechenschrittes die systematischen Instrumentenfehler des Laserscanners sowie die Parameter der äußeren Orientierung zum Stereokamerasystem bestimmt. Ähnlich dem Vorgehen in [90], wo mehrere im Kalibrierfeld verteilte Ebenen unbekannter Lage und Ausrichtung zur inneren Kalibrierung eines TLS verwendet werden, dienen die Ebeneninformationen hier zusätzlich zur Bestimmung der äußeren Orientierung.

Die räumlichen Messpunkte des Stereosystems liefern hierfür Referenzwerte mit übergeordneter Genauigkeit. Dies begründet sich in der in Abschnitt 3.2.1 erwähnten Genauigkeit der absoluten Orientierung von unter 1 mm bis zu einer Entfernung von 7 m und der herstellerseitigen Angabe der Wiederholgenauigkeit einer Laserscanner-Einzelmessung von ca. 10 cm (1σ) [59].

Hierbei werden mit dem Stereosystem und dem Laserscanner sukzessiv mehrere Ebenen gemeinsam vermessen, um anschließend die entsprechenden Ebenenpunkte des Laserscanners über eine Optimierung der Parameter des Fehler- und Orientierungsmodells in die zugehörigen, durch die Stereokamera bestimmten, Referenzebenen zu transformieren. Wie in Abbildung 3.7(a) dargestellt, bestimmt das Stereosystem über korrespondierende Bildpunkte im Stereobildpaar in einem Vorwärtsschnitt die zugehörigen Objektpunkte auf der Ebene E_1 (Vorwärtsschnitt siehe [71] S. 355 ff) und schätzt die Ebenenparameter über einen Ebenenfit [35]. Die Messpunkte des Laserscanners, die ebenfalls E_1 abbilden, werden für die spätere Optimierung selektiert (siehe Abbildung 3.7(b)). Dieses Vorgehen wird für weitere Ebenen, die mit unterschiedlichen Ausrichtungen und Entfernungen den Erfassungsbereich des Laserscanners ausreichend abdecken,

3 Sensorsystem

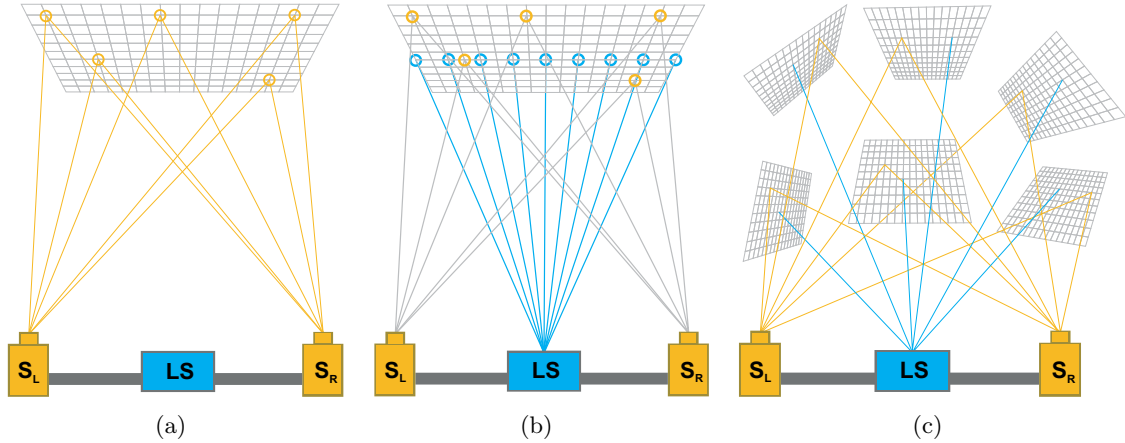


Abbildung 3.7: Bestimmung von Referenzwerten zur Kalibrierung und Orientierung des Laserscanners; Schritt 1: Bestimmung der Referenzebenen durch Stereotriangulation (a); Schritt 2: Auswahl der Ebenen-abbildenden Messpunkte des Laserscanners (b); Wiederholung von Schritt 1 und 2 mit unterschiedlichen Ebenen-Ausrichtungen und Entfernungen (c)

wiederholt, um die anschließende Parameterbestimmung in der Ausgleichung zu stützen.

Vorzugsweise wird die Bestimmung der Referenzebenen im Labor unter Benutzung von photogrammetrischen Kalibriermarken auf den Ebenen durchgeführt, um die Genauigkeit der Bildpunktbestimmung und dementsprechend der Ebenenpunkte zu erhöhen (siehe Abbildung 3.6). Da es zumeist unter den räumlichen Laborbedingungen zu einer Beschränkung der Messentfernung kommt, können die Kalibrierdaten über Messungen im Außenbereich ergänzt werden. Die Bildpunktkorrespondenzen können dabei automatisiert über ein Feature Matching (hier z.B. *SIFT*-Features [70]) ermittelt werden (siehe Abbildung 3.8).

Um der Forderung nach einer höheren Genauigkeit der Referenzwerte nachzukommen, müssen die mit der Entfernung stark zunehmenden Messungenauigkeiten des Stereosystems und fehlerhafte Bildpunkt-Assoziierungen kompensiert werden. Die kürzeste Verbindung der beiden windschief verlaufenden Bildpunktvektoren der korrespondierenden Bildpunkte eines in beiden Bildern identifizierten Merkmals stellt ein Gütemaß zur Objektpunktbestimmung dar. Über einen Grenzwert können so schwache oder fehlerhafte Zuordnungen gefiltert werden. Mit dem RANSAC-Verfahren [33] wird anschließend in einer mehrfachen Iteration der Ausgleichungsrechnung zur Ebenenbestimmung nach dem am besten passenden Parametersatz für die Ebene gesucht und der Einfluss von Datenausreißern reduziert.

Eine über das Stereoverfahren bestimmte Ebene E kann in der impliziten Form durch ihren Normalenvektor \mathbf{N} sowie zweier Ortsvektoren \mathbf{P} und \mathbf{X} auf E mit

$$E : [\mathbf{X} - \mathbf{P}] \mathbf{N} = 0 \quad (3.14)$$

beschrieben werden.

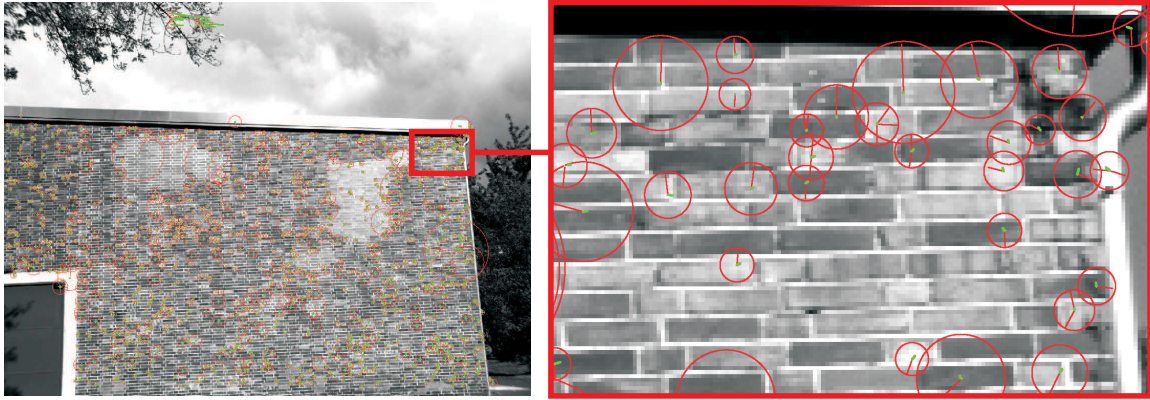


Abbildung 3.8: Ermittelte Features an einer Hauswand zur Bestimmung der Referenzebene (links); vergrößerter Bildausschnitt verdeutlicht die geringe Fehlzuordnung der SIFT-Features durch die kleinen Residuen (grüne Linien) an der Messebene (rechts)

Nach der Überführungsgleichung 3.3 lässt sich ein polarer Laserpunkt L in kartesische Koordinaten mit dem Betrag r und dem Richtungsvektor \mathbf{W} , mit

$$\mathbf{L} = r\mathbf{W} = r \begin{bmatrix} \cos \varphi \cos \theta \\ \sin \varphi \cos \theta \\ \sin \theta \end{bmatrix}. \quad (3.15)$$

darstellen. Über eine Translation \mathbf{T} und eine Rotation \mathbf{R} wird \mathbf{L} durch die Normalenform 3.14 auf E abgebildet:

$$[(\mathbf{R}(r\mathbf{W}) + \mathbf{T}) - \mathbf{P}] \mathbf{N} = 0. \quad (3.16)$$

\mathbf{T} und \mathbf{R} stellen somit die Parameter der äußeren Orientierung des LKS zum SKKS des Stereosystems dar.

Wird Gleichung 3.16 nach der Streckenmessung r_M umgestellt, so erhält man mit

$$r_M = \frac{1}{a} \left(\frac{\mathbf{NP} - \mathbf{NT}}{\mathbf{NRW}} \right) - b \quad (3.17)$$

den funktionalen Zusammenhang zwischen den beobachteten Distanzmessungen und den gesuchten Parametern der äußeren Orientierung und des Fehlermodells über der Gleichung der Referenzebene. Zur korrekten Bestimmung des systematischen Fehleranteils werden mehrere Einzelmessungen (> 20) pro Messpunkt gemittelt, um so den Einfluss der Rauschgrößen u , v und w zu minimieren. Auf die Einführung eines stochastischen Modells, welches die Beobach-

Kalibrierung	Referenzebenen	Mittelwert (mm)	RMS (mm)
Kalibrierung I ohne Fehlermodell	19	1.051	48.9
Kalibrierung I mit Fehlermodell	19	0.006	27.8
Kalibrierung II ohne Fehlermodell	12	-1.294	17.2
Kalibrierung II mit Fehlermodell	12	-0,005	14.9

Tabelle 3.1: Mittlerer Fehler (RMS) und gemittelte Fehlersumme bei der gemeinsamen Bestimmung von systematischem Fehler und äußerer Orientierung des Laserscanners für zwei unterschiedliche Kalibrierungen I und II

gen aufgrund ihrer Streuung gewichtet (wie noch in [51] vorgeschlagen), wird daher verzichtet.

Die so bestimmten Messwerte fließen als Beobachtungen in ein überbestimmtes Gleichungssystem. Durch eine iterative Ausgleichung nach vermittelnden Beobachtungen (siehe hierzu [71] S. 55 ff) werden die gesuchten Parameter nach der Methode der kleinsten Quadrate geschätzt. Die hierfür notwendige Jacobi-Matrix zur Linearisierung des Gleichungssystems wird numerisch bestimmt (siehe [71] S. 72 ff).

Tabelle 3.1 fasst die erzielten Genauigkeiten beispielhaft für zwei Kalibrierungen I und II mit unterschiedlichem Sensoraufbau zusammen. Wird bei der Kalibrierung auf die Verwendung eines Fehlermodells verzichtet und nur die äußere Orientierung zum Stereosystem bestimmt, ist im Gegensatz zur vollständigen Kalibrierung keine Mittelwertfreiheit zu erkennen. Der systematische Fehlereinfluss konnte somit durch die alleinige Orientierungsbestimmung nicht beseitigt werden. Der mittlere Fehler (RMS) bei der Kalibrierung mit einem Fehlermodell ist ebenfalls vergleichsweise geringer. Der insgesamt höhere RMS-Wert bei Kalibrierung I im Vergleich zu Kalibrierung II resultiert aus der Einbeziehung von Referenzebenen, die im Außenbereich bestimmt wurden (siehe Abbildung 3.8).

Das vorgestellte Kalibrierverfahren könnte zudem automatisiert werden, indem günstig gelegene Ebenen automatisch in den Daten der Sensoren bestimmt werden (siehe [90]). Anschließend würden die korrespondierenden Ebenen der beiden Sensoren ermittelt werden (z.B. Ausgleichung nach kleinsten Quadraten). Die Laserpunkte könnten somit automatisch aus den entsprechenden Korrespondenzebenen selektiert und nach dem beschriebenen Verfahren über die zugehörigen Ebenen der Stereodaten der Kalibrierung zugeführt werden.

Bestimmung des zufälligen Fehleranteils

Die Bestimmung der Varianzen der drei polaren Messgrößen r , φ und θ wird in einem Vorschritt der in Abschnitt 3.2.2 beschriebenen systematischen Fehlerbestimmung vollzogen (siehe Abbildung 3.9). Die Schwierigkeit besteht in der Extraktion der Einzelvarianzen aus der Gesamtvarianz $\sigma_{r,\varphi,\theta}^2$ des eindimensionalen Messwerts des Punktabstandes r .

Bei der Ableitung der Varianz σ_r^2 von r aus $\sigma_{r,\varphi,\theta}^2$ müssen die Fehleinflüsse der beiden Po-

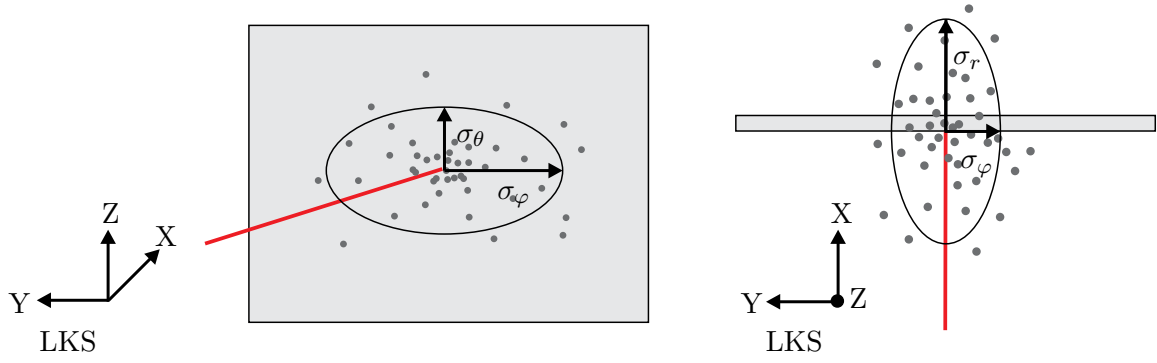


Abbildung 3.9: Modellhafte Streuung der polaren Entfernungsmessung r und der Messwinkel φ und θ an einer orthogonalen Ebene

larwinkel φ und θ abgezogen werden. Die Messentfernung r_l eines orthogonal auf einer Ebene auftreffenden Laserstrahls l wird dabei mit N Einzelmessungen hinreichend oft bestimmt:

$$\sigma_{r_l}^2 = \frac{1}{N-1} \sum_{i=1}^N (r_{l_i} - \bar{r}_l)^2. \quad (3.18)$$

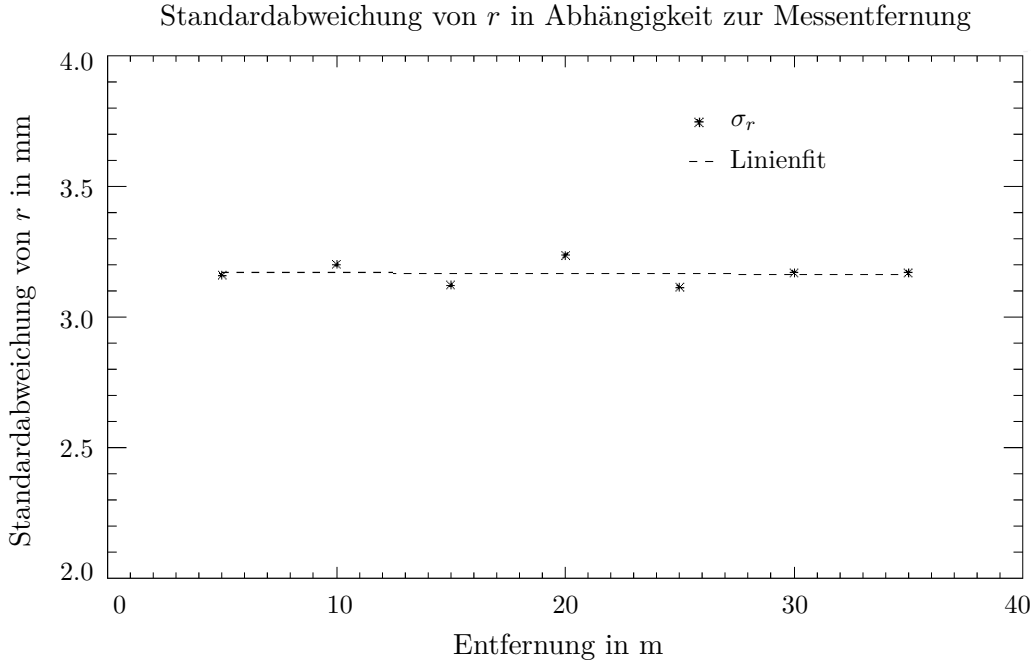
Der Einfachheit halber wird die Varianzschätzung σ_r^2 nur für eine gleichverteilte Untermenge der einzelnen Messrichtungen des Lasersystems vollzogen und σ_r^2 über eine Mittelwertbildung dieser N Einzelvarianzen bestimmt:

$$\sigma_r^2 = \frac{1}{N} \sum_{i=1}^N \sigma_{r_i}^2. \quad (3.19)$$

Wie in Abbildung 3.10 zu erkennen, ist die in Gleichung 3.10 angenommene Unabhängigkeit von u von der Messentfernung zumindest für den Messbereich bis 40 m empirisch bestätigt. Durchschnittlich liegt σ_r hier bei 3.2 mm.

Um die Varianzen σ_φ^2 und σ_θ^2 der beiden Polarwinkel zu bestimmen, müssen diese aus der zusammengesetzten Messstreuung $\sigma_{r,\varphi,\theta}^2$ extrahiert werden. In [40] wird die Standardabweichung des horizontalen Polarwinkels σ_φ hierfür vereinfachend auf die Hälfte der vorgegebenen Winkelauflösung von 0.25° auf $\sigma_\varphi \approx 0.125^\circ$ geschätzt.

Zur empirischen Bestimmung der Polarwinkelvarianzen wurde ein neuer Ansatz entwickelt. Die Einflüsse der Winkelfehler auf den polaren Messwert sollen durch eine Drehung der Reflexionsebene (Referenzebene) an der entsprechenden Drehachse extrahieren werden. Um den Einfluss der Varianz σ_φ^2 auf $\sigma_{r,\varphi,\theta}^2$ zu bestimmen, wird die Referenzebene am Schnittpunkt mit dem Laser um die Z-Achse um einen Winkel α zwischen 0 und $\frac{\pi}{2}$ rotiert. In Abhängigkeit zur Größe der Winkelvarianz σ_φ^2 erhöht sich hierbei die Gesamtvarianz der Messung mit steigendem Ebenenwinkel α und steigender Messentfernung r , da die ideale Strecke r durch den Winkel-

Abbildung 3.10: Unabhängigkeit σ_r von der Messentfernung

fehler von φ und der zur Lasermessung schief stehenden Ebene gestreckt oder gestaucht wird. Die Abbildung 3.11 stellt das simulierte Messverhalten an einer um 45° zur Messrichtung gedrehten Ebene dar. Bei einer Messentfernung von 30 m und normalverteilten Varianzen von $\sigma_\varphi = 0.005 \text{ rad}$ und $\sigma_r = 0.03 \text{ m}$ erkennt man eine schiefe Verteilung der Messpunkte sowie der Entfernungswerte um den idealen Punkt herum.

Die gemessene Strecke \hat{r} lässt sich in Abhängigkeit der Ebenendrehung α , dem Winkelfehler e_φ zwischen dem angenommenen und dem reellen Messwinkel φ und der idealen Entfernung r durch

$$\hat{r} = \begin{cases} \sum_{n=0}^{\infty} \frac{r}{\cos e_\varphi} (\tan \alpha \tan e_\varphi)^n & 0 \leq e_\varphi < \frac{\pi}{2} - \alpha, \\ \frac{r}{\sum_{n=0}^{\infty} \frac{1}{\cos(-e_\varphi)} (\tan(\alpha + e_\varphi) \tan(-e_\varphi))^n} & -\alpha \leq e_\varphi < 0. \end{cases} \quad (3.20)$$

beliebig genau approximieren. Es wird angenommen, dass sowohl die Entfernungsmessung als auch die Winkelbestimmung jeweils einer gaußverteilten Streuung unterliegen und sich die Verteilungen in Form der Messgröße \hat{r} aufsummieren. Für die Faltung $Z = X + Y$ zweier stetiger unabhängiger Zufallsgrößen X und Y gilt bekanntermaßen [97]:

$$\begin{aligned} \mu_Z &= \mu_X + \mu_Y \\ \sigma_Z^2 &= \sigma_X^2 + \sigma_Y^2. \end{aligned} \quad (3.21)$$

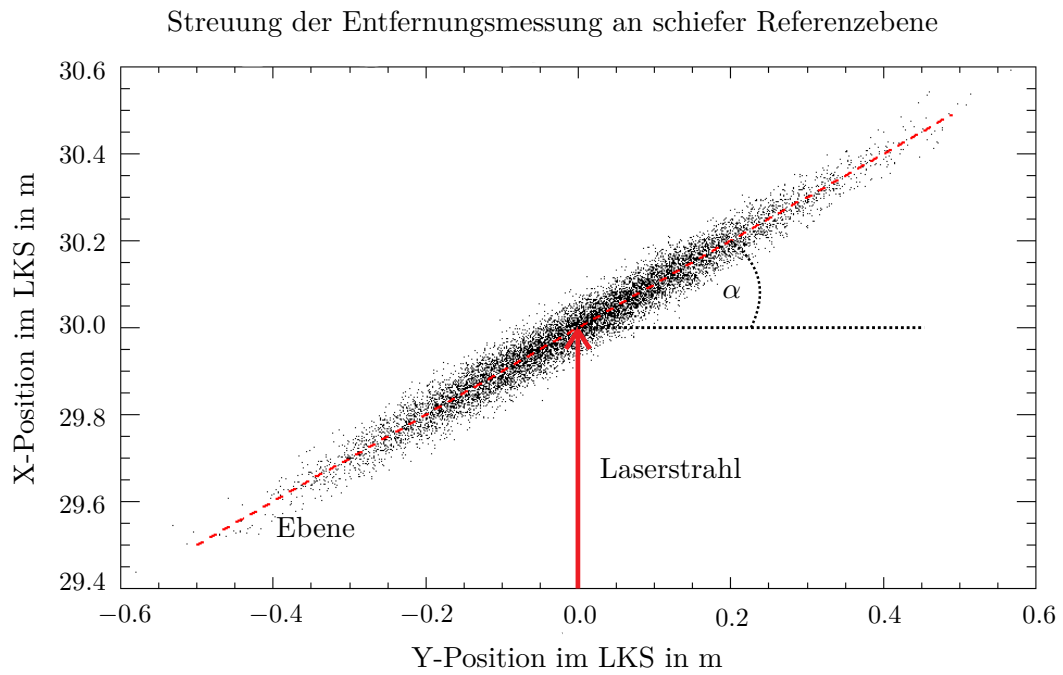


Abbildung 3.11: Simulation der Streuung der Entfernungsmessung an der Referenzebene mit einer Ebenendrehung $\alpha = \frac{\pi}{4}$, einer Messentfernung $r = 30 \text{ m}$, $\sigma_\varphi = 0.005 \text{ rad}$ und $\sigma_r = 0.03 \text{ m}$

3 Sensorsystem

	$\sigma_\varphi = 0.0005$	$\sigma_\varphi = 0.001$	$\sigma_\varphi = 0.005$	$\sigma_\varphi = 0.01$
$\sigma_r = 0.01$	$1.6 \cdot 10^{-7}$	$1.6 \cdot 10^{-6}$	$2.3 \cdot 10^{-5}$	$9.2 \cdot 10^{-5}$
$\sigma_r = 0.02$	$1.3 \cdot 10^{-6}$	$1.3 \cdot 10^{-6}$	$2 \cdot 10^{-5}$	$7.3 \cdot 10^{-5}$
$\sigma_r = 0.03$	$1.4 \cdot 10^{-6}$	$1.9 \cdot 10^{-6}$	$2.4 \cdot 10^{-5}$	$9.7 \cdot 10^{-5}$
$\sigma_r = 0.04$	$2.9 \cdot 10^{-6}$	$3.6 \cdot 10^{-6}$	$2.6 \cdot 10^{-5}$	$1 \cdot 10^{-4}$

Tabelle 3.2: Abweichung der berechneten von der wahren simulierten (horizontalen) Winkelstreuung; Simulation der Streuung der Entfernungsmessung (in m) und Winkelmessung (in rad) mit $\alpha = \frac{\pi}{4}$ und $r = 30$ m in rad

Da die zusammengesetzten Varianzen $\sigma_{r,\varphi}^2$ und $\sigma_{r,\theta}^2$ unter Eliminierung des Einflusses jeweils eines Polarwinkels gemessen werden können und die Varianz der Entfernungsmessung σ_r^2 im Vorfeld bestimmt wurde, lassen sich die einzelnen Varianzen der Winkelmessungen bestimmen durch:

$$\begin{aligned}\tilde{\sigma}_\varphi^2 &= \sigma_{r,\varphi}^2 - \sigma_r^2 \\ \tilde{\sigma}_\theta^2 &= \sigma_{r,\theta}^2 - \sigma_r^2.\end{aligned}\tag{3.22}$$

Die so bestimmte Varianz entspricht der Winkelstreuung entlang der Ebene ohne Einfluss der Streuung der Entfernungsmessung. Zur Überführung der Varianz in die eigentliche polare Winkelvarianz muss die Gleichung 3.20 nach e_φ aufgelöst werden, mit

$$\begin{aligned}\hat{r} &= r + \tilde{\sigma}_\varphi \\ \sigma_\varphi^2 &= e_\varphi^2.\end{aligned}\tag{3.23}$$

Da die Dichtefunktion der Verteilung von \hat{r} nicht symmetrisch, sondern schief ist, werden aus symmetrischen Gründen nur Messwerte $\hat{r} \geq r$ in der Berechnung der Verteilungsparameter verwendet.

Für die Bestimmung der Varianz σ_θ^2 des vertikalen Polarwinkels wird die Referenzebene am Schnittpunkt ausschließlich um die Y-Achse um den Winkel α rotiert, um einen Fehlereinfluss von σ_φ zu vermeiden. Ansonsten gelten analog die Gleichungen 3.20 - 3.23.

Um die Varianzbestimmung zu überprüfen, wurde die Messung modellhaft über die Gleichung 3.20 simuliert. In Tabelle 3.2 ist die Abweichung der verfahrensmäßig bestimmten und der Simulation zugrunde liegenden wahren Standardabweichung der Winkelmessung für unterschiedlich simulierte Streuungsstärken aufgetragen. Die erzielte Genauigkeit zur Bestimmung der Standardabweichung der simulierten Winkelmessung ist dabei in Abhängigkeit zur Winkel- und Entfernungsvarianz mindestens vierstellig.

Bei der Auswertung der realen Kalibriermessungen ergab sich für den Vertikalwinkel θ eine durchschnittliche Standardabweichung mehrerer Laserstrahlrichtungen von 0.0003 rad. Die Streuung des Horizontalwinkels φ war hingegen im Messbereich bis 20 m nicht messbar.

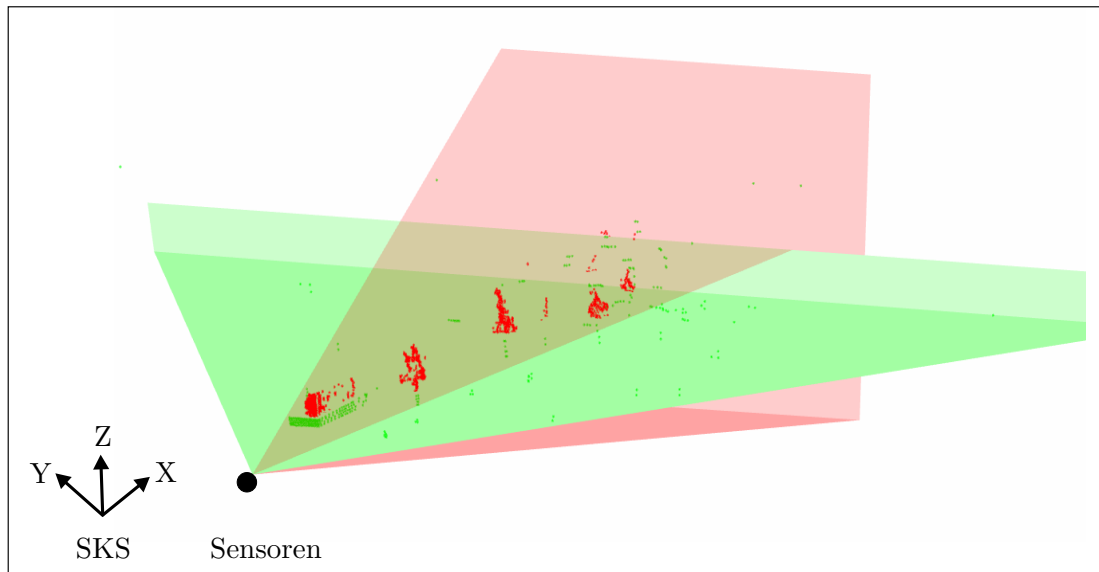


Abbildung 3.12: Darstellung der teilweise überlappenden Sichtbereiche für das Stereosystem (rot) und den Laserscanner (grün); Daten entstammen der Messszene aus Abbildung 3.4

3.2.3 Bestimmung der Sensorsichtbereiche

Nachdem die Sensorrohdaten in ein gemeinsames Koordinatensystem transformiert und von ihren systematischen Fehlern befreit wurden, können die einzelnen Sichtbereiche der Sensoren bestimmt werden. Die Ermittlung erfolgt dynamisch während der gesamten Messung, wobei der Sichtbereich von den Messwerten aufgespannt wird. Eine idealisierte den maximalen Sichtbereich umfassende Bestimmung wird vermieden, um eventuelle Sichtbarkeitseinschränkungen oder sogar Sensorausfälle zu erfassen.

Für das Stereosystem bildet das SKKS im Projektionszentrum der linken Kamera den Ausgangspunkt für den kegelförmigen Sichtbereich, der durch vier Ebenen beschrieben wird. Zusätzlich liefert die verwendete Disparitätensuchtiefe des Stereo-Matchings die minimale Messentfernung des Stereosystems (siehe Gleichung 3.2). Analog erfolgt die Bestimmung für den Laserscanner.

Über die den Sichtbereich beschreibenden Ebenengleichungen in Normalform lässt sich anschließend z.B. die Beobachtbarkeit eines Objektes für einen Sensor überprüfen.

3.3 Ableitung der Datengenauigkeit

Für die weitere Verarbeitung und Fusion der Messdaten ist es notwendig, in Abhängigkeit der räumlichen Lage der Messwerte zum Sensor, eine Genauigkeitsaussage ableiten zu können.

3.3.1 Genauigkeit der Stereodaten

Aus den Gleichungen 3.2 zur Bestimmung der Objektkoordinaten aus einem korrespondierenden Bildpunktpaar im Stereo-Normalfall lassen sich über eine Varianzen-Fortpflanzung aus den Varianzen $\sigma_{x'}^2$ und $\sigma_{y'}^2$ eines Bildpunktes (x', y') des linken Bildes (siehe Gleichungen 3.8 und 3.9) und der Disparitätsbestimmung σ_d^2 die Genauigkeiten der ermittelten Objektkoordinaten berechnen (siehe [65] S. 31):

$$\begin{aligned}\sigma_X &= \frac{X^2}{Bc} \sigma_d \\ \sigma_Y &= \sqrt{\left(\frac{x'}{c} \sigma_X\right)^2 + \left(\frac{X}{c} \sigma_{x'}\right)^2} \\ \sigma_Z &= \sqrt{\left(\frac{y'}{c} \sigma_X\right)^2 + \left(\frac{X}{c} \sigma_{y'}\right)^2}\end{aligned}\tag{3.24}$$

Hierbei werden die Basislänge B und die Kamerakonstante c als fehlerfrei angenommen.

Der Fehler der Disparitätsbestimmung lässt sich nach [71] abschätzen mit

$$\sigma_d = \frac{\sigma_{x'}}{\sqrt{2}}.\tag{3.25}$$

3.3.2 Genauigkeit der Laserdaten

Für die weitere Verarbeitung der Laserscannerdaten in einem einheitlichen kartesischen Koordinatensystem müssen neben den Polarkoordinaten auch die Genauigkeitsangaben in Form der Kovarianzmatrix \mathbf{P} für einen Punkt P mit

$$\mathbf{P}_P = \begin{bmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_\varphi^2 & 0 \\ 0 & 0 & \sigma_\theta^2 \end{bmatrix}\tag{3.26}$$

transformiert werden.

Die Varianzen, die für jeden polaren Messpunkt über die Gleichungen 3.19 und 3.23 zu bestimmen sind, werden nach dem allgemeinen Varianzfortpflanzungsgesetz in das kartesische Koordinatensystem des Laserscanners (LKS) [82] und anschließend in das übergeordnete Koordinatensystem (SKS) überführt. Dafür müssen die Gleichungen aus 3.3 nach den einzelnen Polarkoordinaten abgeleitet werden:

$$\begin{aligned}
 \frac{\partial x}{\partial r} &= \cos \varphi \cos \theta, & \frac{\partial x}{\partial \varphi} &= -r \sin \varphi \cos \theta, & \frac{\partial x}{\partial \theta} &= -r \cos \varphi \sin \theta \\
 \frac{\partial y}{\partial r} &= \sin \varphi \cos \theta, & \frac{\partial y}{\partial \varphi} &= r \cos \varphi \cos \theta, & \frac{\partial y}{\partial \theta} &= -r \sin \varphi \sin \theta \\
 \frac{\partial z}{\partial r} &= \sin \theta, & \frac{\partial z}{\partial \varphi} &= 0, & \frac{\partial z}{\partial \theta} &= r \cos \theta .
 \end{aligned} \tag{3.27}$$

Die Funktionalmatrix A zur Überführung von polaren nach kartesischen Varianzen ergibt sich dann zu:

$$A = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial \theta} \end{bmatrix} \tag{3.28}$$

Analog sind die Gleichungen zur Rotation R_L und Translation T_L eines Vektors L zwischen LKS und SKS mit

$$\begin{aligned}
 x_L &= r_{11}L_x + r_{12}L_y + r_{13}L_z + t_x \\
 y_L &= r_{21}L_x + r_{22}L_y + r_{23}L_z + t_y \\
 z_L &= r_{31}L_x + r_{32}L_y + r_{33}L_z + t_z ,
 \end{aligned} \tag{3.29}$$

nach den Koordinaten von L zur Transformationsmatrix S_L mit

$$S_L = \begin{bmatrix} \frac{\partial x_L}{\partial L_x} & \frac{\partial x_L}{\partial L_y} & \frac{\partial x_L}{\partial L_z} \\ \frac{\partial y_L}{\partial L_x} & \frac{\partial y_L}{\partial L_y} & \frac{\partial y_L}{\partial L_z} \\ \frac{\partial z_L}{\partial L_x} & \frac{\partial z_L}{\partial L_y} & \frac{\partial z_L}{\partial L_z} \end{bmatrix} = R_L \tag{3.30}$$

zu linearisieren, wobei die Jacobi-Matrix S_L wiederum der ursprünglichen Rotationsmatrix R_L entspricht. Der Einfluss der linearen Translation T_L entfällt innerhalb der Varianzfortpflanzung.

Die Kovarianzmatrix P_{P_k} für einen Laserpunkt P_k im kartesischen SKS berechnet sich dann aus der Kovarianzmatrix P_{P_p} des ursprünglichen Laserpunktes P_p in polarer Darstellung mit

$$P_{P_k} = R_L A P_{P_p} A^T R_L^T . \tag{3.31}$$

3.4 Filterung der Punktdaten

Nach der Bereinigung der räumlichen Punktdaten der Sensoren von ihrem systematischen Fehleranteil und der Überführung in ein gemeinsames Koordinatensystem werden durch eine an-

schließende Filterung unnötige oder fehlerhafte Punkte entfernt. Folgende Ziele sollen durch die Filterung der Rohdaten erreicht werden:

- Filterung von Daten außerhalb des Auswertungsbereiches zur Reduktion von Falschdetektionen
- Reduktion von redundanten Informationen zur Verringerung der Rechenzeit der nachfolgenden Prozessierungsschritte
- Bestimmung und Bereinigung von fehlerhaften Daten

3.4.1 Bestimmung des Untergrundes

Durch die enge Kopplung der Fahrbahn mit den sich darauf bewegenden Verkehrsobjekten wird eine fehlerfreie Segmentierung und Objektbestimmung erschwert. Daher ist es sinnvoll den Untergrund modellbasiert zu erfassen und die abbildenden Objektpunkte zu filtern.

Neben einer statischen, im Vorfeld vermessenen linearen Untergrundsmodellierung durch eine Ebene wie in [58], wird die Lage des Untergrundes häufig dynamisch während der Fahrt geschätzt. In [67] wird der Untergrund als längste Gerade im *V-Disparitätenbild* erkannt und ebenfalls über eine Ebene abgebildet. Das V-Disparitätenbild stellt hierbei eine spezielle Darstellungsform der Disparitäten eines Stereobildes dar, das dominante Tiefeninformationen im Bildraum (bspw. eine Straße) über eine Akkumulation aller identischen Disparitätswerte entlang einer Bildzeile als Kanten extrahiert. In [81] wird eine Klothoid-Kurve in die Bodenpunkte approximiert, um den vertikalen Verlauf zu modellieren.

Für diese Arbeit wurde ein halbstatistischer Ansatz gewählt. Im Vorfeld wird ein die Fahrbahn schneidender Quader im SKKS definiert (Quader-Mittelpunkt in $Q_M(15, 0, 0)$). Die Ausmaße des Quaders liegen beispielsweise bei $5\text{ m} \times 5\text{ m} \times 0.5\text{ m}$. Anschließend wird in die quaderumschließenden Objektpunkte eine Untergrundebene geschätzt [35]. Die Güte der Schätzung wird über das *RANSAC*-Verfahren [33] gesteigert. Die Bestimmung der Ebene wird aus Kostengründen in einer geringeren Frequenz, beispielsweise 1 Hz, als der Messtakt durchgeführt.

Analog werden die ins SKS transformierten Laserscannerdaten mit der Untergrundebene gefiltert. Die Fahrbahn wird nicht gesondert bestimmt, da sich hierdurch keine Vorteile für die anschließende Prozessierung ergeben.

Im Gegensatz zum Untergrund wird der Himmel nicht explizit modelliert. Aufgrund der hohen Entfernung von Himmelsobjekten gehen die Disparitäten gegen Null und werden über einen Grenzwert maximaler Erfassungstiefe gefiltert. Fehlerzuordnungen im Stereo-Matching, die aufgrund einer häufig auftretenden homogenen Himmelstextur vorkommen, werden durch die anschließende Reduktion der Stereodaten abgefangen (siehe Abschnitt 3.4.2).

3.4.2 Reduktion der Stereodaten

Um die Reduzierung der Punktwolke möglichst gleichverteilt zu gestalten, wird das Disparitätenbild in quadratische Regionen (Fenster) der Breite n unterteilt. Die Punkte innerhalb eines

Fensters werden an der Position des Fenstermittelpunkts zu einem neuen Punkt P_R zusammengefasst. Hierbei wird der Median \tilde{d} der n^2 der Größe nach geordneten Disparitätswerte aller räumlichen Punkte innerhalb eines Fenster als neue Disparität für P_R ausgewählt:

$$\tilde{d} = \begin{cases} d_{\frac{n^2+1}{2}} & , \text{ für } n \text{ ist ungerade} \\ \frac{1}{2} \left(d_{\frac{n^2}{2}} + d_{\frac{n^2}{2}+1} \right) & , \text{ für } n \text{ ist gerade .} \end{cases} \quad (3.32)$$

Die Positionsgenauigkeit von P_R wird entsprechend Gleichung 3.24 berechnet.

Die durch ein fehlerhaftes Matching entstandenen groben Ausreißer bzw. Randpunkte von Punkthäufungen, die die anschließende Segmentierung verzerren würden, sollen ebenfalls im Zuge der Reduktion gefiltert werden. Hierbei wird eine Tiefenkarte in der Horizontalebene mit gleich großen Regionen erstellt. Um die ein reales Objekt abbildenden Punktverteilungen in der Tiefenkarte lückenlos darzustellen, entsprechen die Zeilen der Tiefenkarte den ansteigenden Disparitätsstufen. Die Spalten bilden eine euklidische Metrik mit konstantem Intervall ab. Alle räumlichen Objektpunkte werden anschließend auf die in der horizontalen X-Y-Ebene liegenden Tiefenkarte abgebildet. Für jede horizontale Region der Tiefenkarte werden die darauf projizierten Objektpunkte mengenmäßig akkumuliert. Durch einen distanzabhängigen Grenzwert wird die Tiefenkarte binarisiert, wodurch Regionen mit zu geringer Belegung als leer deklariert werden. Objektpunkte, die in leeren Regionen liegen, werden anschließend aus der Punktwolke gefiltert (siehe Abbildung 3.13).

Als ein guter Kompromiss zwischen Rauschreduktion und Informationsdichte erwies sich für den dargestellten Sensoraufbau eine Fensterbreite n zwischen sieben und neun Pixeln. Dies entspricht im Objektraum in etwa der horizontalen Auflösung des Laserscanners.

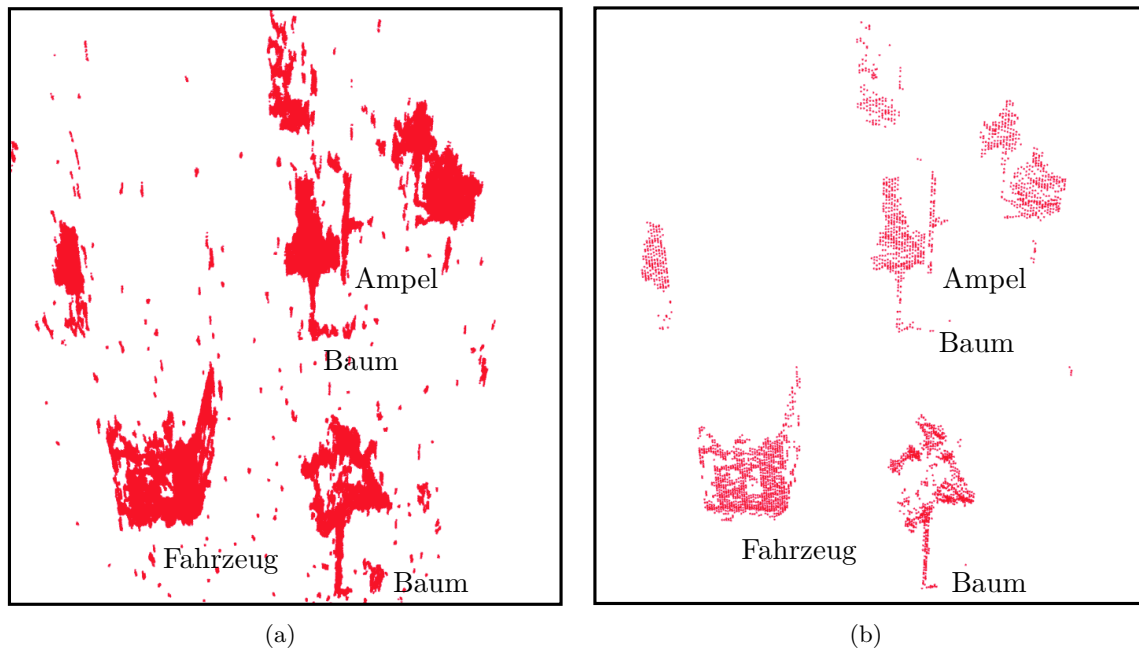


Abbildung 3.13: Stereopunktwolke der Szenerie aus Abbildung 3.4(a) aus der Vogelperspektive vor (a) und nach der Punktreduzierung (b)

3.5 Zusammenfassung

Im Rahmen dieses Kapitels wurden Aufbau und Funktionsweise des entwickelten Referenz-Sensorsystems sowie Verfahren zur Bestimmung und Bereinigung von systematischen und zufälligen Messfehlern und der Datenvorverarbeitung dargestellt. Als wesentliche Neuerungen sind hierbei die Ansätze zur Bestimmung des zufälligen Fehleranteils für den Laserscanner und das Stereokamerasystem sowie die gleichzeitige modellgetriebene Bestimmung der systematischen Fehler und der Koregistrierung des Laserscanners zum Stereosystem beschrieben worden. Auf Grundlage des eingeführten Fehlermodells ist es möglich, für die Punktdaten der Sensoren eine Genauigkeit abzuleiten. Die Datenvorverarbeitung umfasst die Verfahren zur Untergrundmodellierung und Punktfilterung sowie der zweistufigen Reduktion der Stereodaten.

4 Mid Level - Segmentierung und Objektbestimmung

Auf der mittleren Prozessierungsebene werden aus den kalibrierten und vorgefilterten Punkthinformationen eines Sensors räumliche Punkthäufungen in Form von Segmenten extrahiert. In einer weiterführenden Verarbeitung werden die einzelnen Segmente eines Messdurchganges zu Objekthypothesen entsprechend den realen Objekten zusammengefasst.

4.1 Segmentierung

Der Vorgang der Segmentierung bezeichnet im Folgenden die Identifizierung von sinnvoll zusammenhängenden Mengen räumlicher Punkte, die das gleiche reale Objekt abbilden.

Bei bildgebenden Sensoren, wie z.B. einem Mono-Kamerasystem, lassen sich die Segmente über bekannte Punkt- Kanten- oder regionenbasierte Verfahren der Bildverarbeitung nur im 2D-Raum bestimmen. Die Entfernung der Segmente zum Sensor kann nur mit zusätzlichem Priori-Wissen geschätzt werden. In [18] lässt sich z.B. über die Kenntnis des Straßenverlaufs für die darauf befindlichen und extrahierten Objekte die Entfernung abschätzen.

Häufig werden die bildgebenden Sensoren um einen zusätzlichen, räumlich messenden Sensor (z.B. Laser, Radar) ergänzt, sodass die im Bildraum extrahierten Segmente um eine Tiefeninformation des weiteren Sensors ergänzt werden können [113, 1, 31]. Probleme ergeben sich hier jedoch in der dauerhaften Bestimmung der Koregistrierung der Sensoren sowie in der Erkennung von fehlerhaften Detektionen aufgrund der fehlenden Informationsredundanz.

Die räumlichen Punktdaten einer Stereokamera werden klassischerweise auf ein horizontales Gitter projiziert und akkumuliert. Die einzelnen Gitter werden anschließend schwellwertabhängig binarisiert und die zusammenhängenden Bereiche segmentiert [81, 66]. Der Vorteil liegt in einer schnellen Berechnung. Nachteilig wirken sich der Genauigkeitsverlust in der Segmentbestimmung durch die Diskretisierung und der Verlust der Höheninformation aus. Übereinander liegende Objekte (z.B. Baum über einem Fahrzeug) können somit nicht mehr differenziert werden. Um dies zu vermeiden, wird in [58] die aufwendige räumliche Segmentierung von Stereokameradaten durch zwei gekoppelte Segmentierungsschritte im 2D-Raum ersetzt. Hierbei werden die Punkthinformationen zuerst in der klassischen gitterbasierten Draufsicht geclustert. Die so gewonnenen Segmente werden anschließend in der Bildansicht einer Kamera in ihrer Höhe und Breite durch eine Kantenextraktion weiter verfeinert. Zur präzisen Erfassung von sich schnell bewegenden Objekten wird in [36] die Kombination von Tiefen- und Bewegungsinformationen (*Optical Flow*) eines Stereosystems vorgeschlagen, die jedoch eine hohe Berechnungskomplexität aufweist.

Bei aktiv messenden punktgebenden Sensoren wie Laserscanner bzw. Lidar-Systeme mit einer

vergleichsweise geringeren Punktdichte werden die Messpunkte im 2D- oder direkt im 3D-Raum segmentiert. Hierfür wird im einfachsten Falle wie in [13] und [102] die Segmentierung nur durch die Distanzen zwischen den Punkten bestimmt. In [39] wird zusätzlich der Distanzschwellwert zwischen zwei Punkten in Abhängigkeit zur Messentfernung dynamisch bestimmt, um das radiale Messprinzip des Laserscanners zu berücksichtigen. Die Messgenauigkeiten der Punktinformationen werden in [67] mit hinzugezogen. Über eine 2D-Distanzfunktion wird der Abstand zwischen einem Messpunkt und einem Punktecluster in der Draufsicht in Abhängigkeit der Kovarianzen bestimmt. So wird ein Messpunkt einem Cluster zugeordnet, falls sich ihre Konfidenzellipsen überschneiden.

Um das Fusions-Framework vom eigentlichen Sensorverfahren zu trennen, müssen auch die einzelnen Verarbeitungsstufen unabhängig sein. Deshalb stützt sich die Segmentierung nur auf die räumlichen Punktdaten und verzichtet auf die bildhafte Information des Stereosystems. Hierdurch können weitere Sensortypen, die in der Lage sind die Umgebung über Objektpunkte räumlich zu erfassen, in das Prozessierungs-Framework eingebunden werden. Zudem ist für eine Fusion der Punktinformationen eine einheitliche Segmentierung notwendig. Durch eine Erweiterung der Segmentierung in den 3D-Raum ist es möglich die Objektform abzuleiten, um z.B. eine genauere Klassifikation durchzuführen sowie übereinander liegende Objekte zu unterscheiden. In Anlehnung an [39] wird das sensorspezifische Messverhalten (z.B. Messauflösung) sowie die Genauigkeit der Punktinformationen über erweiterte Modelle beschrieben, um die Segmentierung zu steuern.

4.1.1 Segmentierungsmodell

Die Abstraktion des Sensorverfahrens in der Segmentierung soll durch die Modellierung des Segmentierungsverhaltens für jeden Sensortyp erzielt werden. Das Segmentierungsmodell beschreibt für jeden Punkt im Messraum eines Sensors das lokale Segmentierungsverhalten in Abhängigkeit zum Messprinzip und dem Auflösungsverhalten des Sensors sowie der Entfernung des Punktes zum Sensor. Als Segmentierungsbereich wird der Raum um einen Messpunkt bezeichnet, in dem er wahrscheinlich mit einem weiteren, benachbarten Messpunkt des Sensors das gleiche reale Objekt abbildet. Alle Punkte, die dem Segmentierungsbereich eines Punktes angehören, werden somit dem gleichen Segment zugeordnet.

Der durch das jeweilige Messverfahren bedingte, theoretisch minimale Messpunkt Abstand a zweier benachbarter Messpunkte in einer Entfernung d zum Sensor entspricht dem minimalen lokalen Distanzschwellwert zur Segmentierung. Diese Entfernung wird durch die Differenz der Austrittswinkel α zweier benachbarter Punktmessungen am Sensor bestimmt (z.B. Winkel zwischen zwei aufeinander folgenden Laserbeams):

$$a = 2 \tan\left(\frac{\alpha}{2}\right)d = Sd . \quad (4.1)$$

Die lokale Segmentierungsstärke eines Sensors lässt sich durch eine geeignete Wahl des Distanzfaktors S , einem konstanten Offset O und einem Exponenten E beschreiben. Da sich die Messauflösung in den drei Messdimensionen unterscheiden kann (siehe Laserscanner), werden hierfür jeweils individuelle Parametersätze verwendet. Somit lässt sich der lokale Segmentie-

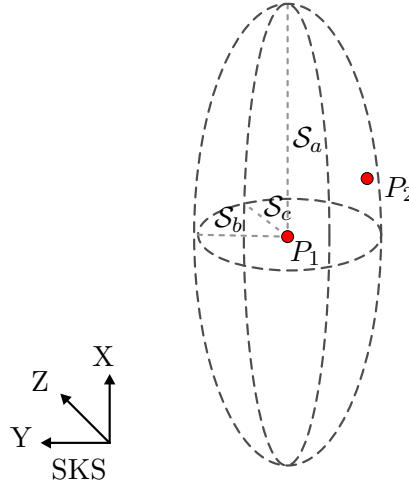


Abbildung 4.1: S-Ellipsoid für einen Messpunkt P_1 , der einen zweiten Messpunkt P_2 einschließt

rungsbereich für einen spezifischen Sensortyp an einem Punkt P als Ellipsoid \mathcal{S}^P (*S-Ellipsoid*) in Abhängigkeit zur Messdistanz d anhand der Halbachsen darstellen:

$$\begin{aligned}\mathcal{S}_a &= O_x + (S_x d)^{E_x} \\ \mathcal{S}_b &= O_y + (S_y d)^{E_y} \\ \mathcal{S}_c &= O_z + (S_z d)^{E_z} .\end{aligned}\tag{4.2}$$

Abbildung 4.1 illustriert den gültigen Segmentierungsraum für einen Messpunkt P_1 , der einen zweiten Messpunkt P_2 umschließt. Dem Modell entsprechend gehören P_1 und P_2 zum gleichen Segment bzw. bilden das gleiche reale Objekt ab.

Neben der Messspezifik der Sensoren wird die Positionsgenauigkeit der Messpunkte in die Segmentierung mit einbezogen. Ein Fehler-Ellipsoid \mathcal{F}^P (*F-Ellipsoid*) gibt dabei den Raum um einen Messpunkt P an, in der sich die wahre Position des Punktes mit einer bestimmten Wahrscheinlichkeit p befindet (siehe Abbildung 4.2(a)). Unter der Annahme, dass die Messung der Punktposition einer Normalverteilung unterliegt, werden die Ellipsoid-Halbachsen folgendermaßen bestimmt:

$$\begin{aligned}\mathcal{F}_a &= \Phi^{-1} \left(\frac{p+1}{2} \right) \sigma_x + \mu_x \\ \mathcal{F}_b &= \Phi^{-1} \left(\frac{p+1}{2} \right) \sigma_y + \mu_y \\ \mathcal{F}_c &= \Phi^{-1} \left(\frac{p+1}{2} \right) \sigma_z + \mu_z, \text{ mit } \mu_x = \mu_y = \mu_z = 0 .\end{aligned}\tag{4.3}$$

Die Standardabweichung σ und der Erwartungswert μ werden in die Standardnormalvertei-

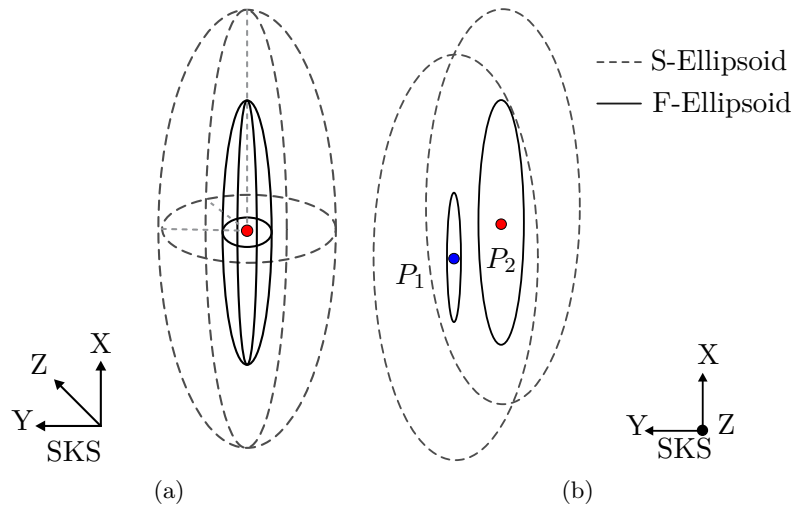


Abbildung 4.2: S- und F-Ellipsoid für einen Messpunkt P_1 (a); gültige Segmentierung zweier Messpunkte in Abhängigkeit von S- und F-Ellipsoid (2D-Ansicht) (b)

lung überführt, um eine Invertierung der Verteilungsfunktion Φ zu vermeiden und auf bereits berechnete Werte in Form einer Lookup-Tabelle zurückzugreifen. Die bestimmten Halbachsenlängen entsprechen demnach einem symmetrischen Streubereich um den Messpunkt mit der Wahrscheinlichkeit p . Da angenommen wird, dass die Messung keinem systematischen Fehleinfluss unterliegt, wird $\mu = 0$ gesetzt.

4.1.2 Segmentierung der Punktinformationen

Im eigentlichen Segmentierungsschritt werden alle Messpunkte paarweise miteinander auf ihre gemeinsame Segmentzugehörigkeit verglichen. Zwei Punkte P_1 und P_2 werden dann einem Segment zugeordnet, wenn, wie in Abbildung 4.2(b) dargestellt, mindestens für einen der beiden Punkte das F-Ellipsoid des einen Punktes im S-Ellipsoid des anderen Punktes vollständig enthalten ist. Somit kann überprüft werden, ob unter der geforderten Wahrscheinlichkeit p die wahren Positionen der beiden Messpunkte im Segmentierungsbereich des jeweils anderen Punktes liegen. Die Überprüfung der Ellipsoid-Zugehörigkeit eines Punktes wird nach [29] durchgeführt.

Die Segmentierung wird somit durch das sensorspezifische Messverhalten und die Punktgenauigkeit über den S-Ellipsoid und den F-Ellipsoid gesteuert (beispielhafte Stereodaten-Segmentierung siehe Abbildung 4.3(b)). Nach der Punktzuordnung an die Segmente wird der minimal umschließende Quader der Segmentpunkte mit paralleler Ausrichtung zu den Koordinatenachsen als Segmentbox bestimmt (beispielhafte segmentierte Sensordaten siehe Abbildung 4.3(a)).

4.1.3 Segmentierungswahrscheinlichkeit

Die Segmentierungswahrscheinlichkeit $P(S_P^X)$ (SW) gibt die Wahrscheinlichkeit für eine Segmentierung S bei der Existenz eines Objektes O am Objektpunkt P für einen Sensor X an. Sie ist

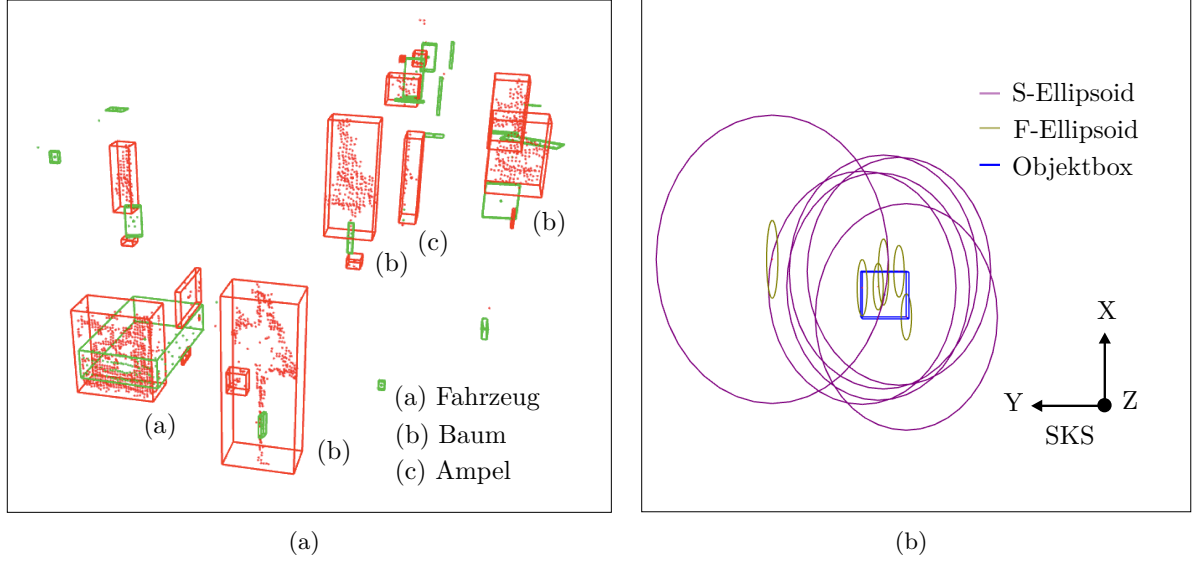


Abbildung 4.3: Beispielhafte Segmentierung von Stereodaten (rot), Laserdaten (grün) in Szene 1 (ohne Fusion) (a) und Darstellung der Segmentierungs- und Fehlerbereiche in einer Segmentdraufsicht (b)

abhängig von der Entfernung und Richtung des Objekts zum Sensor und dem sensorspezifischen Segmentierungs- und Fehlermodell (siehe hierzu Abschnitt 4.1.1).

Die Abschätzung von $P(S_P^X)$ kann für einen beliebigen Punkt P im Messbereich eines Sensors X bestimmt werden. So wird angenommen, dass falls P einem Punkt auf einem beobachtbaren Objekt entspricht, auch die von X messbaren unmittelbaren Nachbarpunkte in der horizontalen Y -Richtung (\bar{P}) und in der vertikalen Z -Richtung ($\bar{\bar{P}}$) von P auf diesem Objekt liegen und von X detektiert werden. Dementsprechend kann eine horizontale SW $P(\bar{S}_P^X)$ zwischen P und \bar{P} und eine vertikale SW $P(\bar{\bar{S}}_P^X)$ zwischen P und $\bar{\bar{P}}$ abgeschätzt werden.

Die horizontale SW wird aus dem kürzesten Abstand \bar{d} zwischen dem F-Ellipsoid $\mathcal{F}^{\bar{P}}$ von \bar{P} und dem S-Ellipsoid \mathcal{S}^P von P und der Standardabweichung $\bar{\sigma}$ von \bar{P} mit

$$P(\bar{S}_P^X) = \begin{cases} \Phi\left(\frac{\bar{d}}{\bar{\sigma}}\right) & , \text{ mit } \bar{d} \geq 0 \\ 1 - \Phi\left(-\frac{\bar{d}}{\bar{\sigma}}\right) & , \text{ mit } \bar{d} < 0 \end{cases} \quad (4.4)$$

geschätzt. Der minimale Abstand \bar{d} zwischen \mathcal{S}^P und $\mathcal{F}^{\bar{P}}$ entspricht dem Quantil für die Normalverteilung des Punktes \bar{P} in der Richtung des kürzesten Abstandes. Ist $\bar{d} = 0$, so berühren sich \mathcal{S}^P und $\mathcal{F}^{\bar{P}}$ und die SW wird mit 50 % bewertet. Für einen negativen Abstand ($\mathcal{F}^{\bar{P}}$ ist nicht mehr komplett in \mathcal{S}^P enthalten) sinkt die Wahrscheinlichkeit unter 50 % und eine Nicht-Segmentierung wird wahrscheinlicher.

Der kürzeste Abstand \bar{d} bestimmt sich vereinfachend aus dem kleinsten der kürzesten Abstände \bar{d}_x , \bar{d}_y , und \bar{d}_z in den drei Koordinatenrichtungen des SKS:

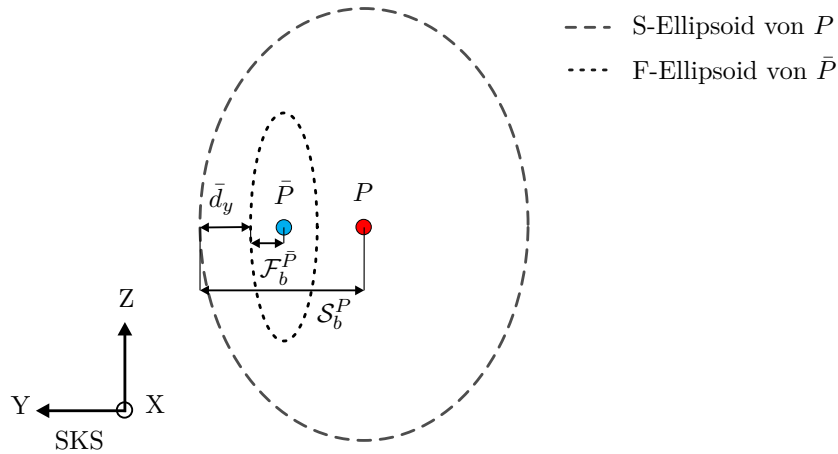


Abbildung 4.4: Darstellung des Abstandes \bar{d}_y in Y-Richtungen zwischen dem S-Ellipsoid \mathcal{S}_b^P von P und dem F-Ellipsoid $\mathcal{F}_b^{\bar{P}}$ von \bar{P} zur Bestimmung der horizontalen SW $P(\bar{S}_P)$ in P

$$\bar{d} = \text{Min}(\bar{d}_x, \bar{d}_y, \bar{d}_z) . \quad (4.5)$$

In Abhängigkeit zur Richtung von \bar{d} wird die Standardabweichung $\bar{\sigma}$ gewählt:

$$\bar{\sigma} = \begin{cases} \sigma_{\bar{P}_x} & , \text{ mit } \bar{d} = \bar{d}_x \\ \sigma_{\bar{P}_y} & , \text{ mit } \bar{d} = \bar{d}_y \\ \sigma_{\bar{P}_z} & , \text{ mit } \bar{d} = \bar{d}_z . \end{cases} \quad (4.6)$$

Die drei Abstände \bar{d}_x , \bar{d}_y , und \bar{d}_z zwischen dem F-Ellipsoid $\mathcal{F}^{\bar{P}}$ des horizontalen Nachbarpunktes \bar{P} sowie dem S-Ellipsoid \mathcal{S}^P von P werden durch

$$\begin{aligned} \bar{d}_x &= \mathcal{S}_a^P - (\bar{P}_x + \mathcal{F}_a^{\bar{P}} - P_x) \\ \bar{d}_y &= \mathcal{S}_b^P - (\bar{P}_y + \mathcal{F}_b^{\bar{P}} - P_y) \\ \bar{d}_z &= \mathcal{S}_c^P - (\bar{P}_z + \mathcal{F}_c^{\bar{P}} - P_z) \end{aligned} \quad (4.7)$$

berechnet (Bestimmung von \bar{d}_y siehe Abbildung 4.4).

Analog wird der vertikale Abstand $\bar{\bar{d}}$ zum F-Ellipsoid $\mathcal{F}^{\bar{\bar{P}}}$ von $\bar{\bar{P}}$ und die vertikale SW $P(\bar{\bar{S}}_P^X)$ ermittelt. Das Minimum von $P(\bar{S}_P^X)$ und $P(\bar{\bar{S}}_P^X)$ bestimmt dann die absolute SW $P(\bar{S}_P^X)$ am Punkt P zum Sensor X :

$$P(S_P^X) = \text{Min}(P(\bar{S}_P^X), P(\bar{\bar{S}}_P^X)) . \quad (4.8)$$

Zur Modellabbildung von Objekten, die nicht achsenparallel zum SKS ausgerichtet sind (z.B. eine zur Messrichtung schiefen Ebene), wird der horizontale Nachbarpunkt \bar{P} zusätzlich in X-Richtung zu P versetzt angenommen. Das Offset in X-Richtung entspricht hierbei dem Abstand von \bar{P} und P in Y-Richtung, um eine durchschnittliche Objektdrehung von 45° zu modellieren.

4.2 Segmentverfeinerung

Anzahl, Form und Ausrichtung der bestimmten Segmente stimmen nicht zwangsläufig mit denen der im Sichtbereich existierenden Objekten überein. In den Schritten der Segmentverfeinerung soll diese Diskrepanz minimiert werden, wodurch der Übergang vom Segment zur Objekthypothese vollzogen wird.

4.2.1 Zusammenfassung von Segmenten

Durch eine konservative Segmentierung, die Objektvereinigungen verhindern soll, können verstärkt Objektfragmentierungen auftreten. Ziel der Segmentzusammenfassung ist es, diese Fragmentierung zu erkennen und die einzelnen Segmente eines Objekts zu vereinen. Hierbei bildet die Überlappung oder ein einseitiger Einschluss zweier Segmentboxen das Kernkriterium für eine Zusammenfassung.

Zusätzlich können auch Segmente vereint werden, deren Mindestabstand unter einem bestimmten Grenzwert d_S liegt. Um eine direkte Abstandsberechnung zu vermeiden, werden die Segmentausmaße um d_S vergrößert und anschließend auf gegenseitige Überschneidung überprüft (siehe Abbildung 4.5).

4.2.2 Bestimmung der Objektausrichtung

Wie in Abbildung 4.7 dargestellt, entspricht die achsenparallel ausgerichtete Segmentbox zumeist nicht der idealen, minimal umschließenden quaderförmigen Hülle des realen Objektes. Diese ist für eine genauere Beschreibung des Objekts, z.B. für eine Klassifizierung über die Objektgröße, notwendig. Zudem sind auch die Positionen der Eckpunkte und des Mittelpunktes, verglichen mit der idealen Objektbox, fehlerbehaftet. Diese Punkte gehen als Referenzpunkte in das anschließende Tracking ein, wodurch der Fehler übernommen wird.

In [20] und [106] wird versucht die minimal umschließende Box zu bestimmen, indem die Randpunkte und der Schwerpunkt des Objektes in einer 2D-Ansicht extrahiert werden. Anschließend wird die Drehung der Hauptachse dahingehend bestimmt, dass die Quadratsumme der orthogonalen Abstände aller Eckpunkte zur Hauptachse minimiert wird. Dieser Ansatz benötigt jedoch eine geschlossene Objektkontur. Bei der Detektion von Verkehrsobjekten ist die gemessene Objektkontur zumeist unterbrochen und unvollständig und das Verfahren kann nicht angewendet werden.

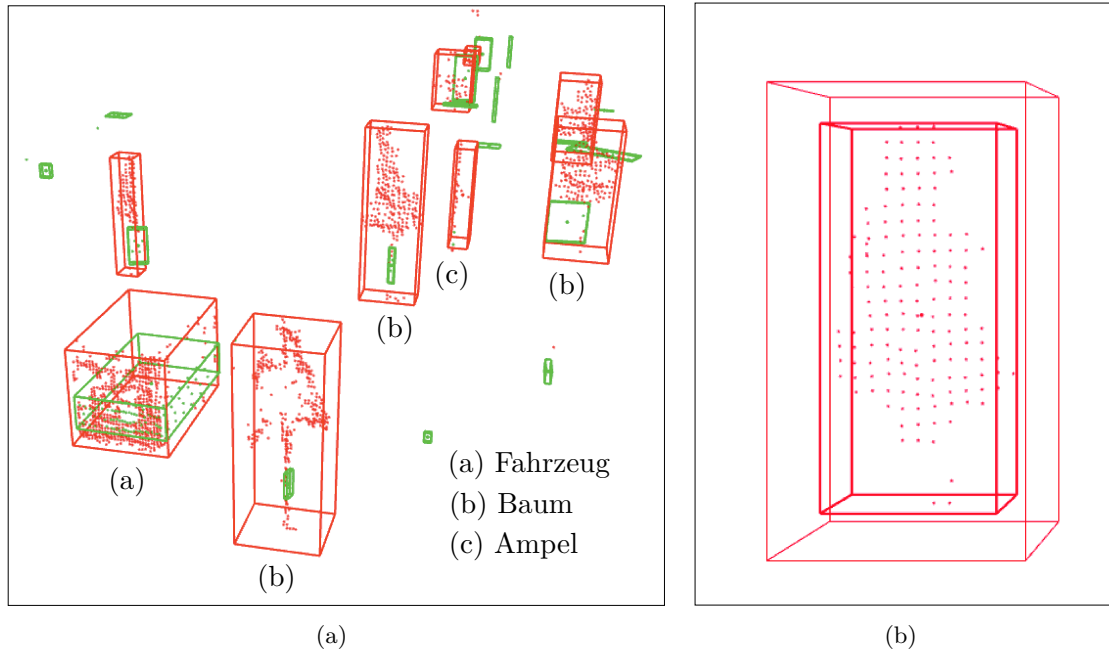


Abbildung 4.5: Segmentboxen der Stereodaten (rot) und Laserdaten (grün) in Szene 1 nach der Zusammenfassung (a); erweiterte Segmentbox zur Zusammenfassung (b)

In [114] werden die Laserscannerpunkte eines detektierten Objektes in drei mögliche Formen (O, I, L) aus der Draufsicht eingeteilt (Abbildung 4.6). Verkehrsobjekte besitzen demnach typischerweise ein Scanprofil der Form *I* oder *L*.

Um die Objektausrichtung im Falle einer L-Form zu ermitteln, werden jeweils der Messpunkt P_{dMin} mit der geringsten Entfernung zum SKS-Ursprung und die Messpunkte mit dem größten (P_{yMax}) und dem kleinsten Wert entlang der Y-Achse (P_{yMin}) im SKS bestimmt. Wie in Abbildung 4.7 dargestellt, wird die Objektform anschließend durch die beiden Strecken zwischen P_{dMin} und P_{yMax} sowie zwischen P_{dMin} und P_{yMin} beschrieben (*Zwei-Linien-Verfahren* [114]). Dabei spannt die Gerade zwischen P_{dMin} und P_{yMax} und der X-Achse den Winkel κ_L und die Gerade zwischen P_{dMin} und P_{yMin} und der X-Achse den Winkel κ_R auf. Der jeweilige Winkel zwischen der X-Achse und einer Geraden wird jeweils in Richtung des Objekts gemessen.

Zur vereinfachten Berechnung werden die beiden Winkel auf eine gemeinsame Richtung (Quadranten) mit

$$\hat{\kappa}_L = \begin{cases} \kappa_L - \frac{\pi}{2} & , \text{ mit } \overline{P_{dMin} P_{yMax}} < \overline{P_{dMin} P_{yMin}} \\ \kappa_L & , \text{ mit } \overline{P_{dMin} P_{yMax}} \geq \overline{P_{dMin} P_{yMin}} \end{cases} \quad (4.9)$$

und

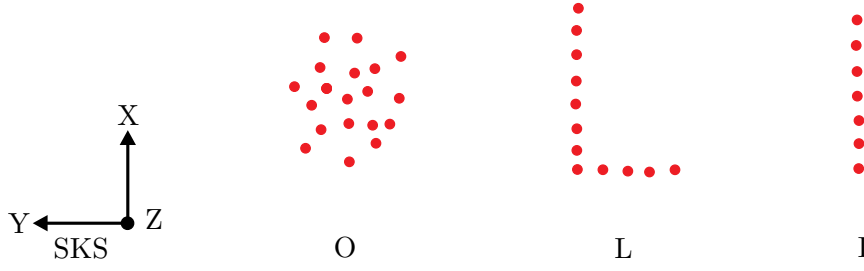


Abbildung 4.6: Einteilung der Punkteverteilung von Objektskans in die Formen O,L und I (nach [114])

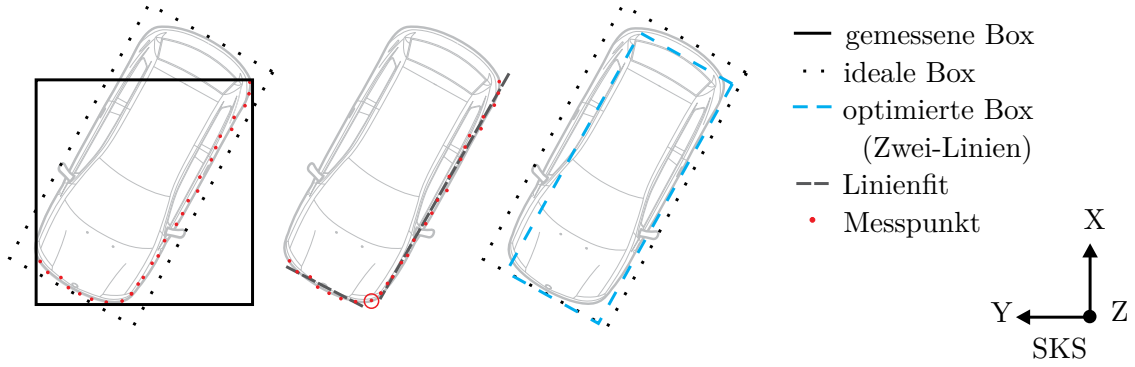


Abbildung 4.7: Unterschiedliche Größe und Ausrichtung zwischen idealer, gemessener Segmentbox und der Segmentbox aus dem Zwei-Linien-Verfahren in der Draufsicht

$$\hat{\kappa}_R = \begin{cases} \kappa_R + \frac{\pi}{2} & , \text{ mit } \overline{P_{d_{Min}} P_{y_{Max}}} \geq \overline{P_{d_{Min}} P_{y_{Min}}} \\ \kappa_R & , \text{ mit } \overline{P_{d_{Min}} P_{y_{Max}}} < \overline{P_{d_{Min}} P_{y_{Min}}} \end{cases} \quad (4.10)$$

normiert. Die Objektdrehung um die Z-Achse κ aus einer Mittlung der beiden normierten Drehungen $\hat{\kappa}_L$ und $\hat{\kappa}_R$ wird wie folgt berechnet:

$$\kappa = \frac{\hat{\kappa}_L + \hat{\kappa}_R}{2} \quad (4.11)$$

Eine Erweiterung dieses Zwei-Linien-Verfahrens wurde für diese Arbeit entwickelt und als *Shaping* bezeichnet. Hierbei werden die Linien nicht allein durch zwei Punkte, sondern durch einen Linienfit über alle Punkte bestimmt [35], um die Anfälligkeit gegen einzelne zu reduzieren. Die Objektdrehung wird als gewichtete Kombination aus den beiden Einzeldrehungen $\hat{\kappa}_L$ und $\hat{\kappa}_R$ der Linien bestimmt. Die Gewichte w_L und w_R entsprechen den jeweiligen RMSE-Werten aus den fehlerhaften Abständen zwischen den Punkten und den beiden Linien:

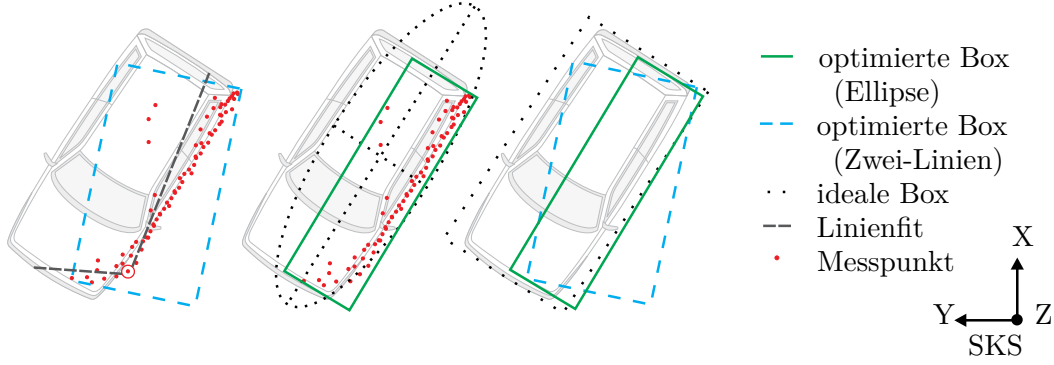


Abbildung 4.8: Vergleich der Segmentbox-Bestimmung zwischen dem Zwei-Linien- und dem Ellipsen-Verfahren

$$\kappa = \hat{\kappa}_L + \frac{w_L}{w_L + w_R} (\hat{\kappa}_R - \hat{\kappa}_L) . \quad (4.12)$$

Die Genauigkeit σ_κ des so bestimmten Winkels κ wird gleich der Abweichung der gemessenen von der idealen Form gesetzt:

$$\sigma_\kappa = \begin{cases} \left| \frac{\Pi}{2} - (|\kappa_L - \kappa_R|) \right| & , \text{L-Form} \\ \left| \Pi - (|\kappa_L - \kappa_R|) \right| & , \text{I-Form} . \end{cases} \quad (4.13)$$

Mit einer steigenden Abweichung des reellen Objektes von der geforderten L- oder I-Form erhöht sich jedoch auch die Ungenauigkeit dieses Ansatzes. Die Abweichungen können zum Beispiel durch Messfehler (z.B. quadratisch ansteigender Fehler beim Stereosystem in X-Richtung), durch einen schrägen Scanwinkel oder komplexe Objektformen, die nicht der geforderten Modellform entsprechen, entstehen.

Um dies zu kompensieren, wird ab einer bestimmten Ungenauigkeit (ca. 5° - 10°) von σ_κ ein robusteres Verfahren angewendet. Über ein Ellipsenfit auf der 2D-Punktmenge des Objektes in der X-Y-Ebene kann die Objektausrichtung selbst bei stärker gestreuten Messwerten hinreichend genau approximiert werden. Die Drehung der Ellipsen-Hauptachse in der X-Y-Ebene dient als Schätzung für den gesuchten Winkel κ (beispielhafte Darstellung in Abbildung 4.8 mit realen Messwerten).

Klassische Ellipse-Fitting-Verfahren sind häufig nicht ellipsenspezifisch und zudem oft rechnerisch anspruchsvoll [44, 92]. In der Arbeit wird ein Ansatz aus [34] verwendet, der über ein direktes Least-Square-Fitting auch für stark gestreute Daten performant und zuverlässig eine Ellipse schätzt. Aus Gründen der Effizienz wird dabei die Anzahl der in das Ellipsen-Fitting eingehenden Messpunkte auf einen bestimmten Maximalwert über eine gleichverteilte Ausdünnung beschränkt.

Analog zum Zwei-Linien-Verfahren wird die Güte der Winkelschätzung für das Ellipse-Fit-

ting-Verfahren bestimmt. Die Güte des Ellipse-Fitting-Verfahrens, welche durch die Distanzen der Objektpunkte zur Ellipse beschrieben wird, hängt nur bedingt mit der Güte der resultierenden Drehwinkelbestimmung zusammen. Daher wird vereinfachend das Seitenverhältnis zwischen Objekttiefe d_x und Objektbreite d_y der gedrehten Objektbox als Gütemaß herangezogen mit

$$\sigma_\kappa = \begin{cases} \frac{d_x \Pi}{4d_y} & , \text{ mit } d_x \leq d_y \\ \frac{d_y \Pi}{4d_x} & , \text{ mit } d_x > d_y . \end{cases} \quad (4.14)$$

Die Gütebestimmung berücksichtigt dabei eine bessere Approximation der Objektdrehung durch das Ellipsen-Fitting für Objekte mit einer langgezogenen L-Form der Messwerte.

Die so bestimmte Objektdrehung κ und deren Genauigkeit σ_κ gehen als Messwerte in die Korrektur des Objektzustands beim Tracking mit ein, um die Zustandsschätzung bei dynamischen Fahrmanövern zu stützen.

Die Segmentzusammenfassung und die Bestimmung der Segmentausrichtung werden solange sukzessiv wiederholt, bis die Menge der Segmente konstant bleibt.

4.2.3 Ableitung der Objektbox

Nach der Bestimmung der zusammenhängenden Segmente und der Schätzung der Objektausrichtung im Raum lässt sich die Objekthypothese ableiten. Hierbei wird das Objektkoordinatensystem (OKS) mit den Achsen parallel zu den entsprechenden Objektseiten modellhaft in den Objektmittelpunkt C gelegt, wobei die X-Achse die Bewegungsrichtung anzeigt. Durch die im Vorschritt geschätzte Objektdrehung κ werden die Objektpunkte um ihren Schwerpunkt vom SKS in das OKS gedreht. Für das somit im SKS achsenparallel ausgerichtete Objekt lässt sich die minimal umschließende Hülle bestimmen. Dabei setzen sich die Eckpunkte (E_1 - E_8) aus den entsprechenden objektaufspannenden minimalen und maximalen Werten der drei Koordinatenrichtungen aller Messpunkte zusammen (siehe Tabelle 4.1). Der Objektmittelpunkt C und die Objektausmaße (d_x , d_y , d_z) leiten sich aus den Eckpunkten ab (siehe Abbildung 4.9). Anschließend werden Eckpunkte und Mittelpunkt um den Schwerpunkt wieder in die ursprüngliche Objektlage zurückgedreht.

Die Objektausmaße leiten sich ebenfalls direkt aus den Abständen der Endpunkte ab. Da die Objekte modellbedingt nur um die Z-Achse rotieren, vereinfacht sich die Bestimmung der Seitenlängen zu:

$$\begin{aligned} d_x &= \sqrt{(E_{1_x} - E_{5_x})^2 + (E_{1_y} - E_{5_y})^2} \\ d_y &= \sqrt{(E_{1_x} - E_{2_x})^2 + (E_{1_y} - E_{2_y})^2} \\ d_z &= E_{3_z} - E_{1_z} . \end{aligned} \quad (4.15)$$

Für die weitere Prozessierung und zur Fusion der Objekte müssen für die einzelnen Eckpunkte, den Mittelpunkt und für die Objektausmaße Genauigkeitsangaben vorliegen. Die Positionsgenauigkeit der Objektecken wird ebenfalls den Genauigkeitsangaben der entsprechenden, das Objekt

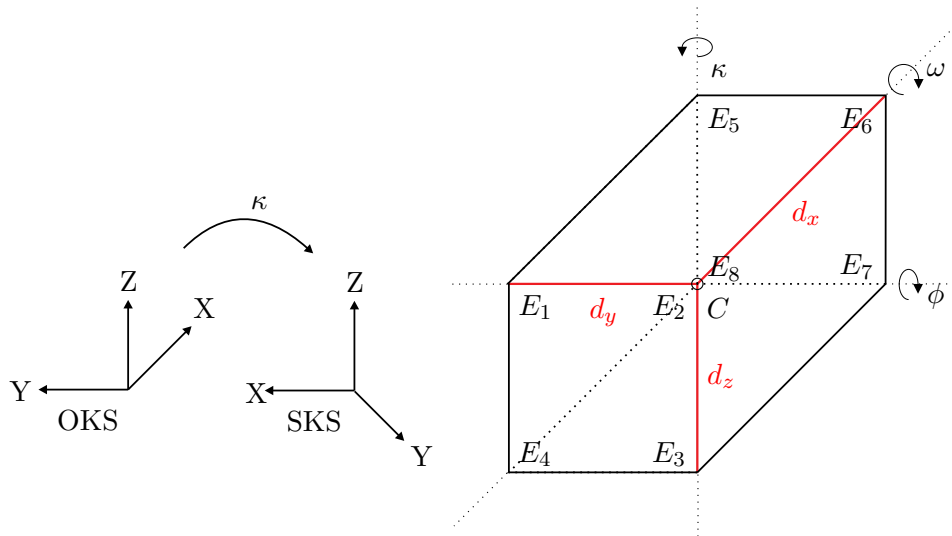


Abbildung 4.9: Eckpunkte ($E_1 - E_8$), Zentrum C , Drehwinkel (ω, ϕ, κ) von OKS nach SKS und Ausmaße (d_x, d_y, d_z) vom Mittelpunkt zur Hülle eines Objekts

Objektpunkt	X-Wert	Y-Wert	Z-Wert
C	$\frac{(X_{min}+X_{max})}{2}$	$\frac{(Y_{min}+Y_{max})}{2}$	$\frac{(Z_{min}+Z_{max})}{2}$
E_1	X_{min}	Y_{max}	Z_{max}
E_2	X_{min}	Y_{min}	Z_{max}
E_3	X_{min}	Y_{min}	Z_{min}
E_4	X_{min}	Y_{max}	Z_{min}
E_5	X_{max}	Y_{max}	Z_{max}
E_6	X_{max}	Y_{min}	Z_{max}
E_7	X_{max}	Y_{min}	Z_{min}
E_8	X_{max}	Y_{max}	Z_{min}

Tabelle 4.1: Bestimmung der Objekteckpunkte $E_1 - E_8$ und des Mittelpunkts C aus den minimalen und maximalen Objektpunktpositionen im SKS

aufspannenden Messpunktkoordinaten entnommen. Die Genauigkeit σ_C der Mittelpunktskoordinaten wird über eine Varianzfortpflanzung der Eckpunkte geschätzt. Hierbei wird die einfache Form des Varianzfortpflanzungsgesetzes [82] angewendet, da die Betrachtung der Kovarianzen zwischen den Messpunkten vernachlässigt wird:

$$\begin{aligned}\sigma_{C_X}^2 &= \frac{\sigma_{X_{min}}^2 + \sigma_{X_{max}}^2}{4} \\ \sigma_{C_Y}^2 &= \frac{\sigma_{Y_{min}}^2 + \sigma_{Y_{max}}^2}{4} \\ \sigma_{C_Z}^2 &= \frac{\sigma_{Z_{min}}^2 + \sigma_{Z_{max}}^2}{4} .\end{aligned}\tag{4.16}$$

Aufgrund der nicht linearen Seitenlängenbestimmung (mit Ausnahme der Länge in Z-Richtung) müssen die Gleichungen 4.15 für die Varianzfortpflanzung linearisiert werden, wobei das totale Differential streng berechnet werden kann mit:

$$\begin{aligned}\sigma_{d_x}^2 &= \left(\frac{\partial d_x}{\partial E_{1x}} \right)^2 \sigma_{E_{1x}}^2 + \left(\frac{\partial d_x}{\partial E_{1y}} \right)^2 \sigma_{E_{1y}}^2 \\ &+ \left(\frac{\partial d_x}{\partial E_{5x}} \right)^2 \sigma_{E_{5x}}^2 + \left(\frac{\partial d_x}{\partial E_{5y}} \right)^2 \sigma_{E_{5y}}^2 \\ &= \left(\frac{E_{1x} - E_{5x}}{\sqrt{(E_{1x} - E_{5x})^2 + (E_{1y} - E_{5y})^2}} \right)^2 \sigma_{E_{1x}}^2 \\ &+ \left(\frac{E_{5x} - E_{1x}}{\sqrt{(E_{1x} - E_{5x})^2 + (E_{1y} - E_{5y})^2}} \right)^2 \sigma_{E_{5x}}^2 \\ &+ \left(\frac{E_{1y} - E_{5y}}{\sqrt{(E_{1x} - E_{5x})^2 + (E_{1y} - E_{5y})^2}} \right)^2 \sigma_{E_{1y}}^2 \\ &+ \left(\frac{E_{5y} - E_{1y}}{\sqrt{(E_{1x} - E_{5x})^2 + (E_{1y} - E_{5y})^2}} \right)^2 \sigma_{E_{5y}}^2\end{aligned}\tag{4.17}$$

$$\begin{aligned}
\sigma_{d_y}^2 &= \left(\frac{\partial d_x}{\partial E_{1_x}} \right)^2 \sigma_{E_{1_x}}^2 + \left(\frac{\partial d_x}{\partial E_{1_y}} \right)^2 \sigma_{E_{1_y}}^2 \\
&+ \left(\frac{\partial d_x}{\partial E_{2_x}} \right)^2 \sigma_{E_{2_x}}^2 + \left(\frac{\partial d_x}{\partial E_{2_y}} \right)^2 \sigma_{E_{2_y}}^2 \\
&= \left(\frac{E_{1_x} - E_{2_x}}{\sqrt{(E_{1_x} - E_{2_x})^2 + (E_{1_y} - E_{2_y})^2}} \right)^2 \sigma_{E_{1_x}}^2 \\
&+ \left(\frac{E_{2_x} - E_{1_x}}{\sqrt{(E_{1_x} - E_{2_x})^2 + (E_{1_y} - E_{2_y})^2}} \right)^2 \sigma_{E_{2_x}}^2 \\
&+ \left(\frac{E_{1_y} - E_{2_y}}{\sqrt{(E_{1_x} - E_{2_x})^2 + (E_{1_y} - E_{2_y})^2}} \right)^2 \sigma_{E_{1_y}}^2 \\
&+ \left(\frac{E_{2_y} - E_{1_y}}{\sqrt{(E_{1_x} - E_{2_x})^2 + (E_{1_y} - E_{2_y})^2}} \right)^2 \sigma_{E_{2_y}}^2 \tag{4.18}
\end{aligned}$$

$$\sigma_{d_z}^2 = \sigma_{E_{3_z}}^2 + \sigma_{E_{1_z}}^2 . \tag{4.19}$$

4.2.4 Segmentierungsfehler

Wie in Abbildung 4.10 dargestellt, weisen die im Zuge der Segmentierung bestimmten Objekthypothesen zumeist einen räumlichen Versatz bzw. Verdrehung sowie abweichende Abmessungen der Objektbox im Bezug zur wahren Objektbox auf (Objektbox-Fehler – *OBF*). Die Ursachen hierfür liegen in der Punktmessgenauigkeit des Sensors, der Unbeobachtbarkeit des gesamten Objekts durch die Sensorperspektive bzw. durch Verdeckungen sowie in einer fehlerhaften oder fehlenden Zuordnung von Messpunkten in der Segmentierung.

Neben dieser Segmentierungsungenauigkeit können im Rahmen der Segmentierung weitere, größere Fehler auftreten. Wie gleichfalls in [101] beschrieben, kann es so zu einer Fragmentierung des realen Objekts in mehrere Objekthypothesen (*Fragmentierungs-Objekt* – *FO*), zu einer Zusammenfassung von mehreren realen Objekten zu einer Objekthypothese (*Zusammenfassungs-Objekt* – *ZO*), zu einer fehlenden Objekthypothese (*Fehldetektions-Objekt* – *FNO*) oder zu einer falschen Objekthypothese (*Falschdetektions-Objekt* – *FPO*) kommen (siehe Abbildung 4.11). Durch die nachfolgenden Verfahren zur Fusion sollen diese Effekte reduziert werden.

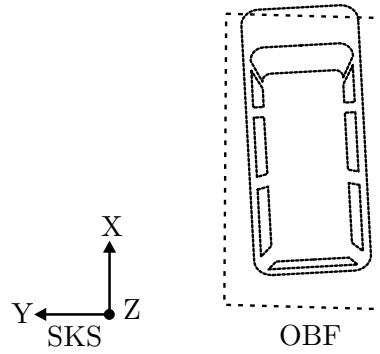


Abbildung 4.10: Abweichung zwischen segmentierter und wahrer Objektbox als Objektboxfehler (OBF)

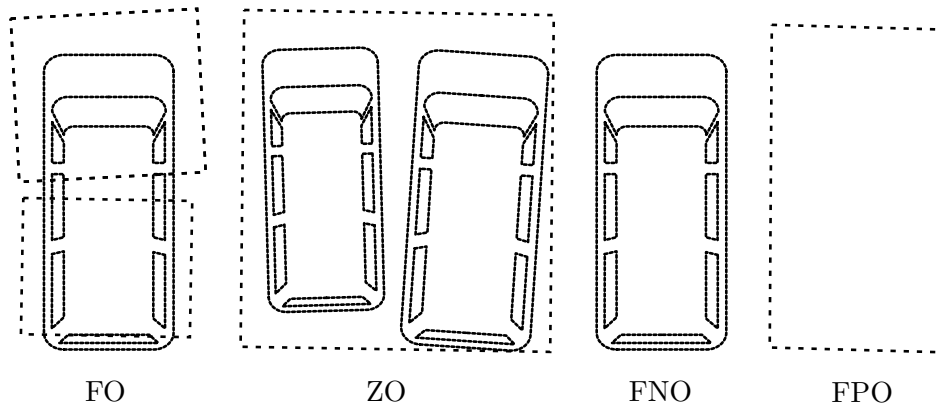


Abbildung 4.11: Segmentierungsfehler: Fragmentierungs-Objekt (FO), Zusammenfassungs-Objekt (ZO), Fehldetektions-Objekt (FNO) und Falschdetektions-Objekt (FPO)

4.3 Zusammenfassung

Die Ableitung von Objektinformationen aus den kalibrierten und vorverarbeiteten Punktdaten der Sensoren wurde im vorangegangenen Abschnitt dargestellt. Um eine ausbalancierte Segmentierung unter Vermeidung von Objektfragmentierung und -zusammenfassung zu erzielen, wurden zur Steuerung der Segmentierung Modelle für die Punktmessgenauigkeit und das sensorspezifische Segmentierungsverhalten integriert. Zudem wurde ein Verfahren zur analytischen Schätzung der Segmentierungswahrscheinlichkeit vorgestellt. Die Ausrichtung eines Objektes, welches sich aus den zusammengehörigen Segmenten ableitet, wird über eine Erweiterung des Zwei-Linien-Verfahrens in Kombination mit einem Schätzverfahren basierend auf Ellipsen-Fitting bestimmt. Neben der Ableitung der Objektparameter und deren Genauigkeiten wurde eine Kategorisierung für die Segmentierungsfehler vorgestellt.

5 High Level - Objektverfolgung

Für eine tiefergehende Beschreibung der detektierten Objekte, insbesondere sich bewegendes Verkehrsobjekte, ist es sinnvoll deren räumlichen Bewegungsverlauf in Form von Tracks auf der oberen Verarbeitungsebene zu ermitteln. Über die Bewegungsinformation ist es möglich das Objekt näher zu charakterisieren, die aktuelle Umfeldsituation zu beschreiben und die zukünftige Bewegung über ein Dynamikmodell kurzfristig vorherzusagen. Im Folgenden werden die Komponenten des hierfür entwickelten *Multi Target Tracking* (MTT) vorgestellt (siehe Abbildung 5.1). In der Objektassoziiierung werden die Objektdetektionen (Messungen) den verfolgten Objektzuständen (Tracks) zugeordnet. In der anschließenden Prädiktion wird über ein Dynamikmodell der aktuelle Objektzustand um einen Zeitschritt in einen zukünftigen Zustand überführt. Dieser prädizierte Objektzustand wird, wenn vorhanden, mit einer Objektdetektion zum gleichen Zeitpunkt im Korrekturschritt abgeglichen und ein optimierter Objektzustand aus beiden Angaben geschätzt. Zudem wird ein Verfahren zur Auswahl sowie zum dynamischen Wechsel des optimalen Tracking-Punktes am Objekt vorgestellt.

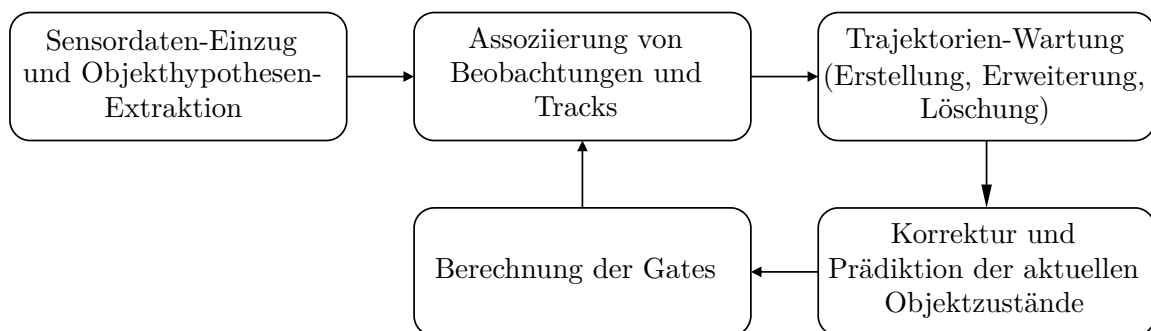


Abbildung 5.1: Grundlegende Funktionsweise der Multi-Objektverfolgung nach dem Prädiktions-Korrektur-Verfahren (nach [10])

5.1 Objektassoziiierung

Wird räumlich nur ein einziges Ziel im *Single Target Tracking* (STT) beobachtet und verfolgt, so ist die Zuordnung zwischen einer aktuellen Objekthypothesen-Prädiktion (Track) und der einer Objekthypothesen-Detektion (Messung) trivial. Sind jedoch mehrere detektierbare Objekte im Beobachtungsbereich vorhanden, die über ein Multi Target Tracking verfolgt werden sollen, so muss die Zuordnung zwischen der Menge der Objekthypothesen-Prädiktionen \mathbb{H}_p und der Menge der Objekthypothesen-Detektionen \mathbb{H}_d durch einen Assoziierungsvorgang bestimmt werden. Die Anzahl der Elemente von \mathbb{H}_d und \mathbb{H}_p müssen dabei nicht zwangsweise übereinstimmen, da die Beobachtbarkeit der Objekte durch den Sensorsichtbereich, die Verdeckung durch andere Objekte sowie weitere Faktoren unbeständig ist. Zudem können sich Fehler in der Prozessierung auf die Mengen \mathbb{H}_d und \mathbb{H}_p auswirken.

Die Umsetzung des MTT in dieser Arbeit ist anhand des Lebenszyklusses eines Objekts von der initialen Detektion bis zum Verlust der Beobachtbarkeit und der Entfernung aus \mathbb{H}_p in Abbildung 5.2 dargestellt.

Wie der Abbildung 5.2 zu entnehmen ist, wird die Assoziation zwischen den Elementen von \mathbb{H}_p und \mathbb{H}_d an mehreren Punkten im Verfolgungsprozess durchgeführt. Somit wirken sich Fehler in der Assoziierung stark in der Güte der Objektverfolgung aus (siehe Abschnitt 5.2.3). Dabei kann es zu Falsch- (*FPZ*) und Fehllassoziationen (*FNZ*) kommen. Die Auswirkungen auf die Objektverfolgung (siehe Abschnitt 5.2.3) sind in Tabelle 5.1 und Abbildung 5.3 dargestellt.

Die Vielzahl an existierenden Verfahren und Abwandlungen zur Zuordnung der Mengen \mathbb{H}_d und \mathbb{H}_p lassen sich grob in drei Gruppen einteilen (nach [10]). Auf die Nennung der *Nearest-Neighbour-Verfahren* wurde aufgrund ihrer schwindenden Bedeutsamkeit verzichtet.

Global Nearest Neighbour (GNN) Hierbei werden die Daten zu einem Messzeitpunkt ausgewertet. Eine Objekthypothesendetektion wird höchstens einer Prädiktion zugeordnet und umgekehrt (*Unique Neighbour Association* – UNA). Dabei werden die Zuordnungen so gewählt, dass die globalen Gesamtkosten (z.B. Distanzen) minimal werden.

All Neighbours Association (ANA) Der All Neighbours Associations Ansatz stellt sich als eine Weiterentwicklung des GNN dar, indem nicht mehr eine wahrscheinlichste Assoziationshypothese, sondern mehrere Hypothesen möglich sind. Dies ist z.B. der Fall, wenn sich

Assoziationsfehler	Auswirkung Objektverfolgung
Falschzuordnung zwischen $h_p \in \mathbb{H}_p$ und $h_d \in \mathbb{H}_d$ (<i>Falsch-Zuordnung</i> – FPZ)	Führt zu Objektwechsel in der Verfolgung und einem fehlerhaften Track (<i>Multi-Objekt-Track</i> – MOT)
Fehlzuordnung zwischen $h_p \in \mathbb{H}_p$ und $h_d \in \mathbb{H}_d$ (<i>Fehl-Zuordnung</i> – FNZ)	Kann u.U. zur Löschung der zugehörigen Objekthypothese aus \mathbb{H}_p und zum Abreißen des Tracks führen (<i>Fehl-Track</i> – FNT)

Tabelle 5.1: Fehler und Auswirkungen in der Assoziierung der Elemente von \mathbb{H}_d und \mathbb{H}_p

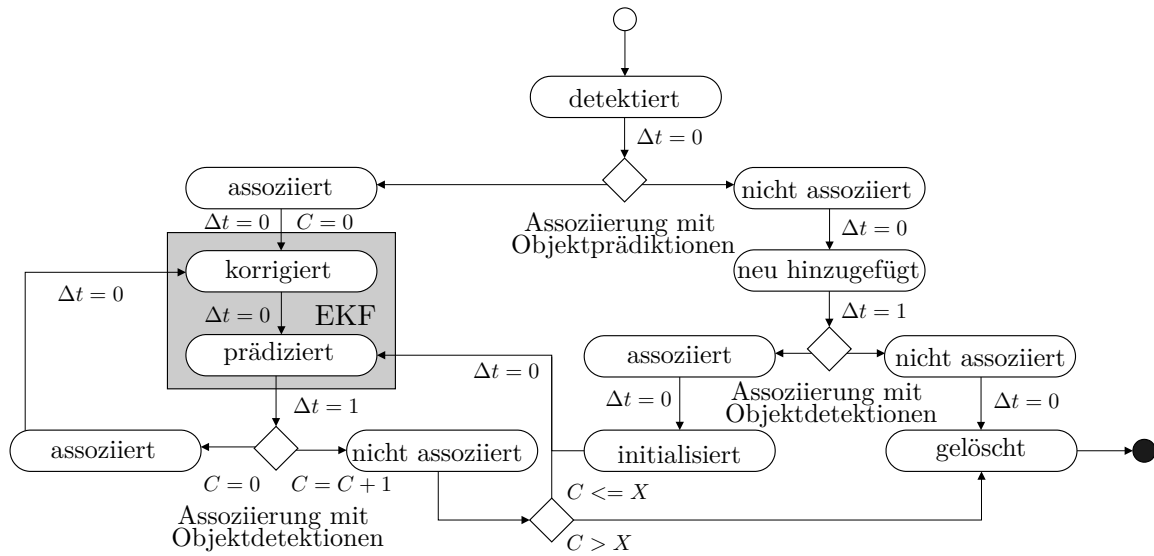


Abbildung 5.2: Lebenszyklus einer Objekthypothese in der Objektverfolgung; die Angaben Δt stehen für die benötigten Zeitschritte zwischen zwei Zuständen im Lebenszyklus, c für die Anzahl aufeinander folgender Zustandsprädiktionen ohne Messwertkorrektur und x für den Schwellwert von c

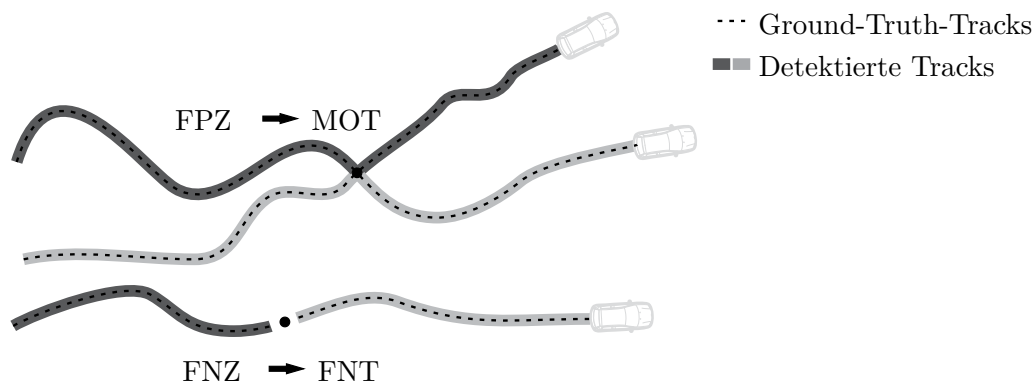


Abbildung 5.3: Darstellung der Auswirkung der Assoziationsfehler Fehl-Zuordnung (FNZ) und Falsch-Zuordnung (FPZ) auf die Objektverfolgung

Verfahren	UNA	ANA	SDL	DDL
GNN	x		x	
PDA/JPDA		x	x	
MHT	x			x

Tabelle 5.2: Einordnung der bekannten Assoziationsverfahren zur Objekthypothesenzuordnung

in dem Gating-Bereich einer Objektprädiktion mehrere detektierte Objekthypothesen (All Neighbours) befinden. Somit können mehrere Beobachtungen gleichzeitig in die Korrektur einer Objektprädiktion eingehen. Erstmals wurde der Ansatz in [6] unter dem Namen *Probabilistic Data Association* (PDA) für STT beschrieben. Zur Erweiterung für ein MTT wurde der Ansatz zum *Joint Probabilistic Data Association* (JPDA) [5] weiter entwickelt. Hieraus sind mittlerweile zahlreiche Modifikationen entstanden. Der in [76] vorgestellte *Joint Integrated Probabilistic Data Association* (JIPDA) integriert beispielsweise zusätzlich ein separates Existenz-Tracking.

Multi Hypotheses Tracking (MHT) Im Gegensatz zu den beiden anderen Gruppen, bei denen die Zuordnungshypothesen innerhalb eines Messschrittes bestimmt werden (*Sequential Decision Logic* – SDL), wird beim MHT die Entscheidung über die wahrscheinlichste Hypothese zu einem späteren Zeitpunkt, wenn mehr Informationen zur Konfliktbereinigung der Mehrfach-Zuordnung vorhanden sind, aufgelöst (*Deferred Decision Logic* – DDL).

Wie in [10] angedeutet und in Tabelle 5.2 dargestellt, lassen sich diese drei Gruppen nach der Eindeutigkeit der Zuordnung (UNA oder ANA) sowie der Anzahl der einfließenden Messschritte (SDL oder DDL) klassifizieren.

Aus Gründen der Echtzeitfähigkeit der Objektverfolgung wurde für die Arbeit ein Assoziationsverfahren entwickelt, dass bei einer steigenden Anzahl an Beobachtungen keinen exponentiellen Berechnungsaufwand, wie bei den bekannten globalen Verfahren, benötigt. Zudem sollte die Fehl- und Falschzuordnungsrate auf einem vergleichbaren Niveau liegen. Zusätzlich zur klassischen Objekt-zu-Track-Zuordnung sollte das Verfahren auch auf die Objekt-zu-Objekt- und die Track-zu-Track-Zuordnung anwendbar sein, um es auf den verschiedenen Fusionsebenen anwenden zu können.

Ausgangspunkt bilden hierfür die Arbeiten von [98] und [85], in denen Merkmale aus zwei Bildern über ihre paarweisen Eigenschaften Abstand d und Ähnlichkeit C zueinander assoziiert werden. Über eine Gauß-Gewichtung dieser Eigenschaften wird für jeweils zwei Merkmale i und j ein initialer Assoziationswert G_{ij} in der Gleichung

$$G_{ij} = [e^{-(C_{ij}-1)^2/2\gamma^2}] e^{-d_{ij}/2\lambda^2} \quad (5.1)$$

bestimmt. Da die beiden einzelnen Hauptterme in Gleichung 5.1 einen Wertebereich von $(0, 1]$ besitzen, liefert auch das Produkt G_{ij} Werte im gleichen Bereich und kann als Assoziationswahrscheinlichkeit zweier Merkmale interpretiert werden.

Über die Streuungsparameter γ und λ werden die zu erwartenden Standardabweichungen für die Assoziierungsparameter C und d festgelegt und somit der Einfluss dieser Eigenschaften auf G bestimmt.

Im nächsten Schritt wird die so erstellte Assoziationsmatrix G zwischen den Merkmalen zweier Bilder einer Singulärwertzerlegung (SVD) unterzogen

$$G = TDU^T . \quad (5.2)$$

Die so erzeugte Matrix D wird in eine neue Matrix E überführt, wobei alle Diagonalelemente auf eins gesetzt werden. Die Lösungsmatrix F ergibt sich somit aus

$$F = TEU^T \quad (5.3)$$

und hat die bemerkenswerte Eigenschaft, dass sie gute Assoziierungspaarungen verstärkt und schlechte abschwächt.

Zwei Merkmale i und j werden als zugeordnet angesehen, wenn das Matricelement F_{ij} das Zeilen- und Spaltenmaximum für Zeile i und Spalte j bildet. Somit ist sichergestellt, dass nur eindeutige Zuordnungen bestimmt werden. Des Weiteren wird nach [85] durch F eine Zuordnungskonfiguration mit einer globalen Distanzminimierung erzeugt, da die Spur von $F^T G$ maximiert wird. Das Verfahren lässt sich somit in die Klasse der GNN-Verfahren einordnen.

Um das Verfahren auf die Assoziierung von Objektdetektionen und verfolgten Objektprädiktionen anzuwenden, werden die Merkmale als Objekthypothesen interpretiert. Die Eigenschaft d entspricht dementsprechend dem Abstand zwischen den Referenzpunkten einer Objekthypothesenpaarung. Sinnvollerweise sollte hierbei die Streuung λ kleiner als der Gating-Radius g gewählt werden (siehe Abschnitt 5.1.1). In der Arbeit hat sich ein Verhältnis von $\lambda = g/3$ als praktikabel erwiesen.

Im Falle der Assoziierung von Objekthypothesen basierend auf Stereokameradaten wird die Ähnlichkeitseigenschaft C aus der mittleren Bildpunktintensität der objektbeschreibenden Bildpunkte bestimmt. Hierbei wird C nach der Anzahl der möglichen Intensitäten normiert. Eine denkbar schärfere Objektcharakterisierung durch den Vergleich von Intensitätshistogrammen der objektbeschreibenden Bildpunkte hat sich in der Praxis nicht ergeben.

Analog wird C für Objekthypothesen, die auf Laserdaten basieren, aus der Mittelung der normierten Intensitätswerte (Echo-Breite) der objektbeschreibenden Laserpunkte bestimmt.

5.1.1 Metrik der Objektdistanzen

Neben dem Ähnlichkeitsmaß geht der Abstand d zwischen einer Objekthypothesenprädiktion und einer Objekthypothesendetektion zur Bestimmung der Assoziation mit ein. Die Distanzfunktion d bestimmt dabei den Abstand der Zustandsgrößen zweier Objekte aus der Menge der Objekte \mathbb{O} als einen positiven, eindimensionalen Wert mit

$$d : \mathbb{O} \times \mathbb{O} \longrightarrow \mathbb{R}_0^+ . \quad (5.4)$$

Die Abbildung d entspricht genau dann einer Metrik auf \mathbb{O} , wenn folgende drei Axiome erfüllt sind:

1. Selbstidentität: $\forall O \in \mathbb{O} : d(O, O) = 0$
2. Symmetrie: $\forall O_1 \neq O_2 \in \mathbb{O} : d(O_1, O_2) = d(O_2, O_1)$
3. Dreiecksungleichung: $\forall O_1, O_2, O_3 \in \mathbb{O} : d(O_1, O_3) \leq d(O_1, O_2) + d(O_2, O_3) .$

Als bekannter Vertreter misst die *euklidische* Abstandsfunktion d_E als Spezialfall der *Minkowski-Metrik* den Abstand zweier n -dimensionalen Punkte P_1 und P_2 im euklidischen Raum \mathbb{R}^n , wobei

$$d_E(P_1, P_2) = \sqrt{(P_1 - P_2)^T (P_1 - P_2)} . \quad (5.5)$$

Um neben den Punktpositionen die Unsicherheiten der Messgrößen in die Distanzbestimmung mit einzubeziehen, bietet sich die Mahalanobis-Distanz d_M [72] in der Form

$$d_M^{P_1}(P_1, P_2) = \sqrt{(P_1 - P_2)^T \mathbf{S}_{P_1}^{-1} (P_1 - P_2)} \quad (5.6)$$

an. Durch die Verwendung der invertierten Kovarianzmatrix \mathbf{S} verringert sich die Mahalanobis-Distanz d_M bei steigender Unsicherheit und erhöht sich bei steigender Sicherheit zwischen zwei Punkten bei einer konstanten euklidischen Distanz. Im zwei-dimensionalen Fall kann dies anschaulich durch eine Drehung um 90° der Hauptachsen zwischen Fehlerellipse und Distanzellipse aufgrund der Invertierung von \mathbf{S} dargestellt werden. Somit bilden im zwei-dimensionalen Fall alle Punkte mit gleichem euklidischen Abstand zu einem Punkt P_1 über d_E einen Kreis und über d_M eine Ellipse um P_1 .

Für den Spezialfall, dass die Kovarianzmatrix \mathbf{S} der Einheitsmatrix gleicht, entspricht die Mahalanobis-Distanz der Euklidischen Distanz. Abbildung 5.4 begründet die Wahl der Metrik d_M für die Assoziierung von Punkten. Obwohl die Punkte P_2 und P_3 die gleiche euklidische Distanz zu P_1 besitzen, weist die Metrik d_M dem unsicheren P_2 eine kleinere Distanz als dem sicheren P_3 zu, da die Wahrscheinlichkeit, dass P_2 an der Stelle von P_1 liegen könnte, höher ist als bei P_3 .

Wie in [18] ebenfalls bemerkt wird, kann es jedoch zu einer Verzerrung der Objektassoziation durch die Benutzung der Mahalanobis-Distanz kommen, falls eine Objekthypothese eine unverhältnismäßig hohe Positionsvarianz besitzt, wodurch ihre Abstände zu den zu assoziierenden Objekten sehr klein wird. Dies ist z.B. bei einer Objekthypothese möglich, die seit mehreren Zeitschritten nicht korrigiert, sondern nur prädiiziert wurde, wodurch die Unsicherheit des Objektzustandes ansteigt.

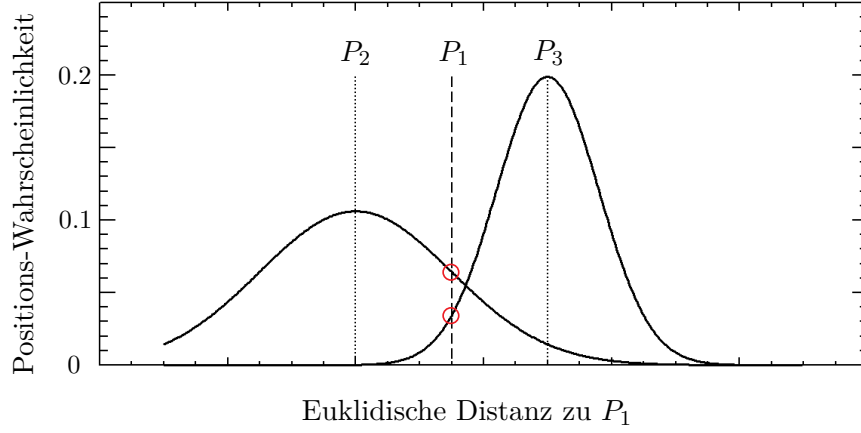


Abbildung 5.4: Obwohl die Punkte P_2 und P_3 den gleichen euklidischen Abstand zu P_1 haben, besitzt der unsichere Punkt P_2 eine höhere Wahrscheinlichkeit an der Stelle von P_1 zu liegen, was durch die Mahalanobis-Distanz abgebildet wird

Um dies zu vermeiden, wird in [18] vorgeschlagen, die maximale Distanz der beiden einzelnen Mahalanobis-Distanzen zweier Punkte als eigentliche Assoziierungsdistanz zu wählen. In einer leichten Variation dessen wird hier in der Berechnung der Assoziierungsdistanz $d_{\bar{M}}$ die Kovarianzmatrix \mathbf{S} als Mittel der beiden einzelnen Objekt-Kovarianzmatrizen \mathbf{P}_{P_1} und \mathbf{P}_{P_2} gebildet, um so die Varianzen beider Objekte zu bewerten:

$$d_{\bar{M}}(P_1, P_2) = \sqrt{(P_1 - P_2)^T \mathbf{S}_{P_1 P_2}^{-1} (P_1 - P_2)}, \text{ mit} \quad (5.7)$$

$$\mathbf{S}_{P_1 P_2} = \frac{1}{2}(\mathbf{P}_{P_1} + \mathbf{P}_{P_2}). \quad (5.8)$$

Da die Assoziierung jeweils zwischen zwei Objekten aus den Mengen \mathbb{H}_p und \mathbb{H}_d vollzogen wird, bestimmt sich \mathbf{S} durch das Mittel der beobachtbaren Prozessvarianzen $\mathbf{H}\mathbf{P}\mathbf{H}^T$ und der Messvarianzen \mathbf{R} zu

$$\mathbf{S}_{P_1 P_2} = \frac{1}{2}(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}). \quad (5.9)$$

Zur weiteren Schärfung der Distanzbestimmung zweier Objekte fließen neben den Referenzpunktkoordinaten P_1 und P_2 zusätzlich die drei Objektdimensionen d_x, d_y, d_z in die Distanzfunktion mit ein, wodurch sich ein 6-dimensionaler Zustandsvektor O ergibt. Die Matrizen \mathbf{P} , \mathbf{H} , \mathbf{R} und \mathbf{S} erweitern sich entsprechend. Die Gleichung der Distanzfunktion lautet demnach:

$$d_{\bar{M}}(O_1, O_2) = \sqrt{(O_1 - O_2)^T \mathbf{S}_{O_1 O_2}^{-1} (O_1 - O_2)}. \quad (5.10)$$

Da die Abbildung $d_{\bar{M}}$ einer Metrik d_M mit einer neu bestimmten Kovarianz S entspricht, wird davon ausgegangen, dass auch $d_{\bar{M}}$ selbst die Axiome einer Metrik (Selbstidentität, Symmetrie und Dreiecksungleichung) erfüllt.

Durch das *Gating* wird versucht, aufgrund eines zu großen Abstandes sehr unwahrscheinliche Assoziierungspaare im Vorfeld zu erkennen, um so den Assoziierungsprozess auf die Menge der potentiellen Paarungen auf \mathbb{H}_p und \mathbb{H}_d zu beschränken, um somit fehlerhafte Zuordnungen und den Berechnungsaufwand zu reduzieren (siehe Abbildung 5.1 zum MTT). Die Gating-Distanz g wird dabei im Vorfeld manuell gesetzt und als Grenzwert für eine zulässige Assoziierungsdistanz $d_{\bar{M}}$ zweier Punkte verwendet. Hierbei wird für die initiale Assoziierung zwischen zwei zeitlich versetzten Detektionen aufgrund der fehlenden Prädiktion ein höherer Wert für g als für Assoziierungen zwischen Prädiktionen und Detektionen des gleichen Zeitschritts angesetzt.

5.2 Objekttracking

Der Begriff Tracking bezieht sich hier auf die Schätzung und Vorhersage der Parameter des Bewegungsmodells eines verfolgten Objekts auf der Grundlage von Messwerten. Klassische Verfahren beruhen auf einem Least-Square Schätzer (LSE), der über die Werte einer zeitlichen Messreihe die Zustandsparameter unter einer gemeinsamen Optimierung nach kleinsten Quadraten bestimmt. In einer rekursiven Erweiterung des LSE kann der zuletzt bestimmte Zustand zu jedem neuen Messzeitpunkt mit aktuellen Werten abgeglichen (korrigiert) werden. Durch eine zusätzliche Modellierung des Zufalls zu den Parametern des Objektzustands ergibt sich der lineare Kalman Filter (KF) sowie dessen zahlreichen Erweiterungen, wie z.B. den hier verwendeten nichtlinearen erweiterten KF (EKF) [112].

Komplexere Ansätze verfolgen, zumeist auf Grundlage des KF, mehrere Bewegungsmodelle für ein Objekt parallel, um diese entsprechend den einzelnen Schätzgenauigkeiten gewichtet zu kombinieren oder auszuwählen (z.B. *Interacting Multiple Model* – IMM [11, 10]).

Auf eine komplexere Bewegungsmodellierung wie beim IMM wurde in dieser Arbeit jedoch verzichtet, um die Fehleinflüsse beim Tracking gering zu halten und somit die Vergleichbarkeit der getrackten Objektdaten der unterschiedlichen Fusionsebenen zu wahren.

Ein Großteil der zu beobachteten Objekte stellen in dieser Arbeit Fahrzeuge dar. Daher wurde ein nichtlineares, zeitdiskretes Systemmodell zur Schätzung eines Systemzustandes $\tilde{\mathbf{x}} \in \mathbb{R}^n$ ausgewählt, dass in der Lage ist, kurvige Bewegungsverläufe zu approximieren.

Die Zustandsschätzung wird durch die Erweiterung des linearen Kalman Filters über eine Linearisierung an den Systemzuständen realisiert [112]. Der aktuelle Systemzustand $\tilde{\mathbf{x}}_k$ zum Zeitpunkt k wird dabei mittels der nichtlinearen Funktion f , einem bekannten Eingangsvektor \mathbf{u}_k und dem Prozessrauschen ν_k einen Zeitschritt weiter in $\tilde{\mathbf{x}}_{k+1}$ überführt

$$\tilde{\mathbf{x}}_{k+1} = f(\tilde{\mathbf{x}}_k, \mathbf{u}_k, \nu_k) . \quad (5.11)$$

Über die nichtlineare Funktion h wird der Systemzustand $\tilde{\mathbf{x}}$ zur Messung $\tilde{\mathbf{z}} \in \mathbb{R}^m$ und zum Messrauschen ξ im Beobachtungsmodell in Beziehung gesetzt

$$\tilde{Z}_k = h(\tilde{X}_k, \xi_k) . \quad (5.12)$$

Der Zustand eines zu verfolgenden Objektes wird durch die linear beschreibbaren Größen der Referenzpunktposition x , y , z , der Geschwindigkeit am Referenzpunkt v und den Objektausmaßen d_x , d_y , d_z beschrieben. Die Modellierung von gleichförmigen Kreisbewegungen im Systemmodell soll durch die Zustandserweiterung um die Drehwinkel am Referenzpunkt ψ_R und des gesamten Objektes ψ_O sowie den entsprechenden Drehwinkelgeschwindigkeiten ω_R und ω_O möglich werden:

$$\tilde{X}_k = \begin{bmatrix} x_k & y_k & z_k & d_{x_k} & d_{y_k} & d_{z_k} & v_k & \psi_{R_k} & \omega_{R_k} & \psi_{O_k} & \omega_{O_k} \end{bmatrix}^T . \quad (5.13)$$

In [12] wird hierbei von einem polaren Systemmodell gesprochen. Die Beschleunigungen von v_k , ω_R und ω_O werden nicht explizit modelliert, sondern über das Prozessrauschen implizit abgebildet (siehe Gleichung 5.18). Die getrennte Modellierung von Bewegungsrichtung und Drehwinkelgeschwindigkeit für Gesamtobjekt und Referenzpunkt sind nötig, da sie auch in der Realität unterschiedlich sind. Bei der Korrektur der Referenzpunkte mit dem gemessenen Winkel der Gesamtobjektausrichtung kam es folglich zu einer Zustandsverfälschung, die eine Zustandstransformation beim dynamischen Referenzpunktwechsel erschwert (siehe Abschnitt 5.3.2).

5.2.1 Prädiktion des Objektzustandes

Die Transformation des Systemzustands um einen Zeitschritt zum Zeitpunkt $k + 1$ wird in den Systemgleichungen mit

$$\hat{X}_{k+1}^- = f(\hat{X}_k, U_k, 0) = \begin{bmatrix} x_k + \cos(\psi_k)v_k\Delta t \\ y_k + \sin(\psi_k)v_k\Delta t \\ z_k \\ d_{x_k} \\ d_{y_k} \\ d_{z_k} \\ v_k \\ \psi_{R_k} + \omega_{R_k}\Delta t \\ \omega_{R_k} \\ \psi_{O_k} + \omega_{O_k}\Delta t \\ \omega_{O_k} \end{bmatrix} \quad (5.14)$$

vollzogen, wodurch der *a posteriori* Zustand \tilde{X}_k zum Zeitpunkt k in den *a priori* Zustand \tilde{X}_{k+1}^- zum Zeitpunkt $k + 1$ prädiziert wird. Da die jeweiligen Realisierungen für das Prozess- und Messrauschen (ν , ξ) zu den unterschiedlichen Zeitpunkten nicht bekannt sind, werden sie in den Systemgleichungen und den Beobachtungsgleichungen (siehe 5.21) vernachlässigt und \tilde{X} durch \hat{X} und \tilde{Z} durch \hat{Z} näherungsweise approximiert.

5 High Level - Objektverfolgung

Zusätzlich zur Zustandspropagierung wird im Prädiktionsschritt des EKF die aktuelle *a posteriori* Kovarianzmatrix des Zustands $\hat{\mathbf{P}}_k$ um einen Zeitschritt zur *a priori* Kovarianzmatrix $\hat{\mathbf{P}}_{k+1}^-$ zum Zeitpunkt $k + 1$ prädiziert:

$$\hat{\mathbf{P}}_{k+1}^- = \mathbf{A}_{k+1} \hat{\mathbf{P}}_k \mathbf{A}_{k+1}^T + \mathbf{W}_{k+1} \mathbf{Q} \mathbf{W}_{k+1}^T . \quad (5.15)$$

Neben der Jacobi-Matrix \mathbf{A} der partiellen Ableitungen von f nach \hat{x} mit

$$\mathbf{A}_{k+1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -\sin(\psi_k)v_k\Delta t & \cos(\psi_k)\Delta t & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \cos(\psi_k)v_k\Delta t & \sin(\psi_k)\Delta t & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.16)$$

sowie der Jacobi-Matrix \mathbf{W} der partiellen Ableitungen von f nach ν mit

$$\mathbf{W}_{k+1} = \begin{bmatrix} \frac{1}{2} \cos(\psi_k)(\Delta t)^2 & 0 & 0 \\ \frac{1}{2} \sin(\psi_k)(\Delta t)^2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ \Delta t & 0 & 0 \\ 0 & \frac{1}{2}(\Delta t)^2 & 0 \\ 0 & \Delta t & 0 \\ 0 & \frac{1}{2}(\Delta t)^2 & 0 \\ 0 & \Delta t & 0 \end{bmatrix} \quad (5.17)$$

gehen auch die Kovarianzen des Systemmodells mit

$$\mathbf{Q} = \begin{bmatrix} \sigma_{a_v}^2 & 0 & 0 \\ 0 & \sigma_{a_\omega}^2 & 0 \\ 0 & 0 & \sigma_d^2 \end{bmatrix} \quad (5.18)$$

in die *a priori* Bestimmung der Zustandskovarianzen mit ein. Hierbei wird das Prozessrauschen

\mathbf{Q} durch die geschätzten Varianzen $\sigma_{a_v}^2$ der Bewegungsbeschleunigung, der Drehwinkelbeschleunigung $\sigma_{a_\omega}^2$ sowie der gemeinsamen Varianz σ_d^2 zur Messung der Objektboxdimensionen d_x , d_y , d_z und der Z-Position des Referenzpunktes z definiert.

5.2.2 Korrektur des Objektzustandes

Die durch die Prädiktion bestimmten *a priori* Schätzungen des Systemzustandes $\hat{\mathbf{X}}_{k+1}^-$ und der zugehörigen Kovarianzen $\hat{\mathbf{P}}_{k+1}^-$ zum Zeitpunkt $k+1$ werden im Korrekturschritt des EKF durch die Messwerte \mathbf{Z} und deren Kovarianzen \mathbf{R} über die Gewichtsmatrix \mathbf{K} zu den *a posteriori* Schätzungen des Systemzustandes mit

$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_k^- + \mathbf{K}_k(\hat{\mathbf{Z}}_k - \mathbf{H}\hat{\mathbf{X}}_k^-) \quad (5.19)$$

und den *a posteriori* Kovarianzen des Zustands durch

$$\hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\hat{\mathbf{P}}_k^- \quad (5.20)$$

überführt.

Als Messgrößen gehen die Referenzpunktpositionen x , y und z am Objekt, die durch das Shaping geschätzte Objektdrehung κ und die Objektboxdimensionen d_x , d_y und d_z ein. Die Beobachtungsgleichungen ergeben sich somit zum Zeitpunkt k aus

$$\hat{\mathbf{Z}}_k = h(\hat{\mathbf{X}}_k, 0) = \mathbf{H}\hat{\mathbf{X}}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \hat{x}_k. \quad (5.21)$$

Die Objektausrichtung einer quaderförmigen Objektdetektion O' lässt sich ohne Kenntnis der Bewegungsrichtung neben der gemessenen Drehrichtung κ noch durch drei weitere mögliche Drehrichtungen beschreiben. Diese ergeben sich durch die Addition der Vielfachen von $\frac{\Pi}{2}$ zu κ mit

$$\kappa_n = \kappa + n\frac{\Pi}{2}, \text{ mit } n \in \{0, 1, 2, 3\}. \quad (5.22)$$

Um einen Messwert für die Bewegungsrichtung ψ_O der zugeordneten Objektprädiktion O'' abzuleiten, wird die Objektausrichtung κ_n von O' mit dem geringsten Winkelabstand zur Bewegungsrichtung ψ_O von O'' gewählt. Die so gemessene Objektausrichtung κ (siehe Gleichung 4.12) schätzt den Drehwinkel der Objektbewegung ψ_O über dem Mittelpunkt der Hinterachse.

5 High Level - Objektverfolgung

Da dieser Punkt aufgrund von Verdeckung am Objekt nicht exakt bestimmbar ist, kann der Messwert κ nicht fehlerfrei auf die Bewegungsrichtung ψ_R eines anderen Referenzpunktes überführt werden. Aus diesen Grund wird, wie bereits zuvor erwähnt, die Objektausrichtung ψ_O im Systemmodell vom Referenzpunkt-Tracking getrennt verfolgt.

Die Gewichtungsmatrix K (häufig auch als *Gain* bezeichnet) wird klassischerweise über das Verhältnis der *a priori* Zustands-Kovarianzen \hat{P}^- und Mess-Kovarianzen R bestimmt, um die *a posteriori* Schätzung der Zustands-Kovarianzen \hat{P} zu minimieren (eine tiefergehende Begründung ist z.B. in [75] zu finden):

$$K_k = \hat{P}_k^- H^T (H \hat{P}_k^- H^T + V R_k V^T)^{-1} . \quad (5.23)$$

Die Jacobi-Matrix V der partiellen Ableitungen der Beobachtungsgleichungen h nach ξ entspricht der Einheitsmatrix und kann dementsprechend vernachlässigt werden.

5.2.3 Fehler in der Objektverfolgung

In den Arbeiten [101] und [100] wird in zwei grundsätzliche Verfolgungsfehler unterschieden. Zum einen führt ein Objektwechsel innerhalb der Verfolgung zu einem fehlerhaften Track (hier als *Multi-Objekt-Track* – MOT bezeichnet). Dies kann z.B. für eine spätere Situationsbeschreibung anhand der Track-Verläufe zu falschen Schlüssen führen. Durch den MOT-Effekt kann es zudem zum Abreißen des ursprünglich verfolgten Tracks und zum Verlust des Bewegungsverlaufes kommen.

Des Weiteren kommt es zum Abreißen eines Tracks, falls keine aktuellen Objektdetektionen mehr zugeordnet werden können (hier als *Fehl-Track* – FNT bezeichnet).

Zusätzlich werden das Auftreten eines Tracks ohne reelle Objektentsprechung (*Falsch-Track* – FPT) sowie, analog zum Objektboxfehler, der Schätzfehler des Objektzustandes (*Objektzustandsfehler* – OZF) als weitere Fehler in der Objektverfolgung definiert. Abbildung 5.5 stellt die Fehler-Typen beispielhaft dar.

Die Ursachen für ein FPT liegen auf der Segmentierungsebene in einem Fragmentierungs-Objekt (FO) oder Falsch-Objekt (FPO). Ein FNT kann dementsprechend durch ein Zusammenfassungs-Objekt (ZO), Fehl-Objekt (FNO) bzw. in einer Fehlzurordnung einer Objekthypothesendetektion (FNZ) hervorgerufen werden. Ein MOT wird durch eine Falschzurordnung (FPZ) hervorgerufen (siehe auch Abbildung 6.1).

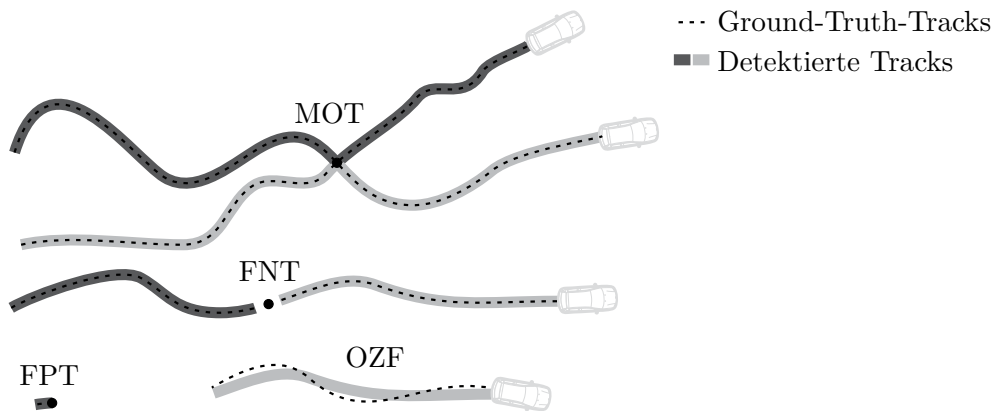


Abbildung 5.5: Kategorisierung der Tracking-Fehler in Multi-Objekt-Track (MOT), Fehl-Track (FNT), Falsch-Track (FPT) und Objektzustandsfehler (OZF)

5.3 Objektbeobachtbarkeit

Die komplette oder partielle Beobachtbarkeit eines Objektes hängt im Wesentlichen von der Begrenzung des Sichtbereiches des Sensors sowie möglichen Verdeckungen durch andere Objekte oder von der teilweisen Eigenverdeckung durch das Zielobjekt ab.

Des Weiteren kann die Objektbeobachtbarkeit vom Messverfahren des Sensors abhängig sein. Durch eine hohe Entfernung oder eine zu geringe Objektgröße kann das Objekt bedingt durch die Messauflösung des Sensors in einen detektionsfreien Bereich fallen. Zudem können sich auch die Objekteigenschaften, wie Texturierung oder Reflektionsverhalten oder Umweltfaktoren wie Regen, Nebel oder Dunkelheit negativ auf die Detektionsfähigkeit des Sensorverfahrens auswirken.

Der folgende Abschnitt beschäftigt sich mit der Bestimmung der kompletten oder partiellen Beobachtbarkeit eines zu verfolgenden Objektes infolge von Verdeckungen oder Sichtbereichseinschränkungen sowie den negativen Folgen und Verbesserungsvorschlägen für das Tracking.

5.3.1 Referenzpunktbestimmung

Die Festlegung des optimalen Referenzpunktes zur Objektverfolgung wird vor dem eigentlichen Tracking-Prozess im Rahmen der Objektbestimmung durchgeführt. Für eine optimale Schätzung des Bewegungsverlaufes eines Objektes müssen sich die prädiizierte Position aus dem Tracking und die dazu assoziierte gemessene Position auf die gleiche physische Referenz am Objekt beziehen.

Die Verwendung des geometrischen Schwerpunkts der Objektpunkte eignet sich dabei nur bedingt für größere Objekte wie Fahrzeuge, da dieser stark von der Fahrzeugausrichtung zum Sensor und der resultierenden Messkontur abhängt.

Der Mittelpunkt der Objektbox als Referenz, wie in [52] dargestellt, ist ebenfalls nur unzureichend genau, da die wahre Objektbox und somit der wahre Mittelpunkt nur selten hinreichend genau bestimmt werden können. Abbildung 5.6 verdeutlicht dies am Beispiel eines in den Sen-

5 High Level - Objektverfolgung

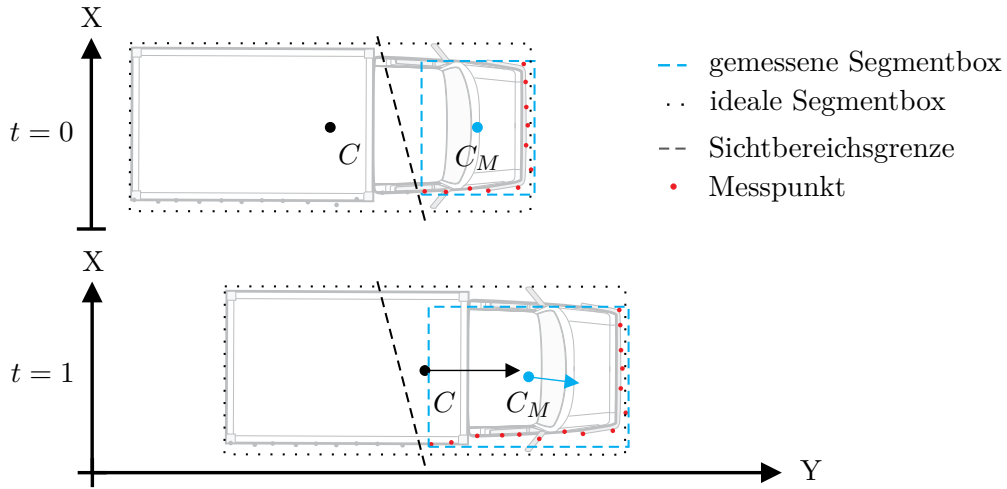


Abbildung 5.6: Abweichung der gemessenen und idealen Objektbewegung über eine Bestimmung durch den idealen C und dem gemessenen Mittelpunkten C_M

sorsichtbereich fahrenden Fahrzeugs. Hierbei ist zu erkennen, dass die Länge und Richtung des nach einem Zeitschritt gemessenen Bewegungsvektors (blau dargestellt) für den Objektmittelpunkt (C_M) im Vergleich zum idealen Bewegungsvektor (schwarz dargestellt) des wahren Objektmittelpunktes (C) deutlich abweicht. Dieser systematische Fehler überträgt sich in einer fehlerhaft geschätzten Dynamik im Tracking-Prozess. Die resultierende fehlerhafte Prädiktion der Objektbewegung kann so zu einer Verzerrung der Situationsanalyse (z.B. fehlerhafte Kollisionserkennung) oder zu einer falschen oder fehlenden Objektassoziiierung führen.

In [41] wird davon ausgegangen, dass sich die 2D-Kontur der Messpunkte eines Fahrzeuges durch eine oder zwei zueinander senkrechte Linien abstrahieren lässt. Ausrichtung und Länge der Linien werden mit Zusatzwissen zur Objektlänge und -breite und der Bewegungsrichtung aus der Objekthistorie korrigiert und dienen der Bestimmung der optimalen Referenzpunkte. Nachteilig können sich hierbei jedoch Fehler im Vorabwissen, die zu einer kontinuierlichen Wiederholung des Fehlers führen können (Rückkopplungsfehler) sowie nur partiell beobachtbare Objektkanten auf die Ergebnisse auswirken.

Ein geeigneter Referenzpunkt sollte daher beobachtbar und in anschließenden Objektaufnahmen eindeutig wiedererkennbar sein. Demnach bietet es sich an, sich auf die sichtbaren Eckpunkte eines Objektes zu beziehen. In [18] werden ebenfalls neben den Kantenmittelpunkten, die Eckpunkte einer 2-dimensionalen Objektform (Rechteck) als mögliche Referenzpunkte definiert und je nach Lage und Beobachtbarkeit zum Sensor ausgewählt. Mögliche Abweichungen der gemessenen von den wahren Referenzpunktpositionen etwa durch Verdeckungen anderer Objekte, des Zielobjektes oder dem Sensorsichtbereich werden nicht behandelt.

Die Beobachtbarkeit der wahren Objektecken E_1-E_8 (siehe Abbildung 4.9) werden in dieser Arbeit zum einen über die Lage des Objektes im Sichtbereich des Sensors und zum anderen über mögliche Verdeckungen durch andere Objekte oder des Objekts selbst bestimmt. Da sich die Objekte modellhaft nur um die Z-Achse drehen, werden die jeweils übereinander liegenden

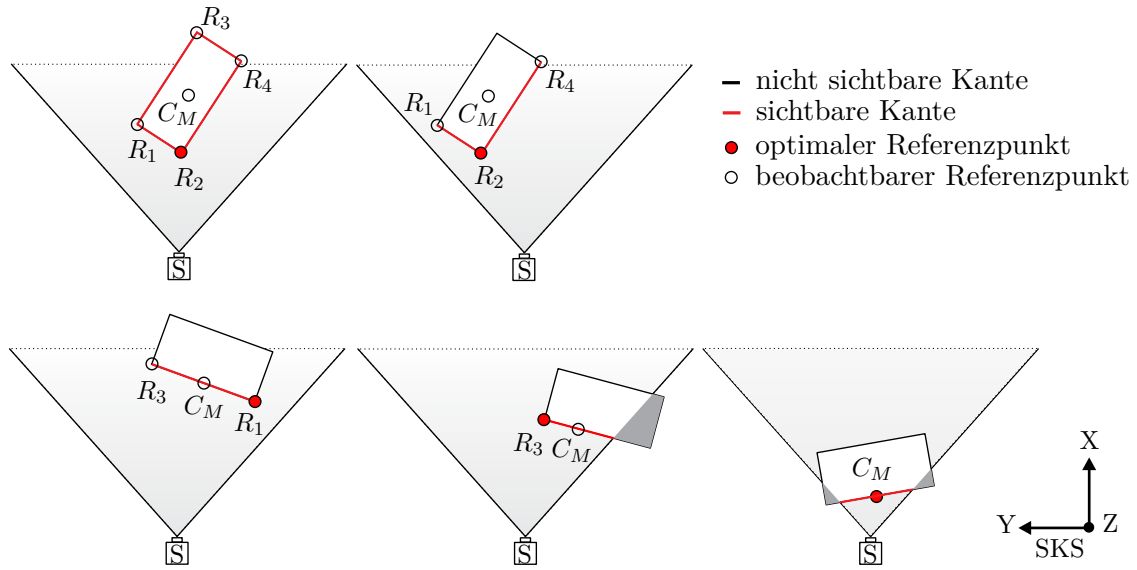


Abbildung 5.7: Bestimmung des optimalen Objekt-Referenzpunktes unter Beachtung des Sichtbarkeitsbereiches des Sensors

Eckpunkte, ohne Betrachtung der Höheninformation, zu den vier zwei-dimensionalen Referenzpunkten R_1 (E_1, E_4), R_2 (E_2, E_3), R_3 (E_5, E_8) und R_4 (E_6, E_7) zusammengefasst. In Abhängigkeit zum Sichtbarkeitsbereich eines Sensors ergeben sich dabei, wie in Abbildung 5.7 aufgeführt, fünf verschiedene Fälle für die Eckpunktbeobachtbarkeit und die daraus beobachtbaren und optimalen Referenzpunkte.

Der Fall der kompletten Beobachtbarkeit aller vier Referenzecken ist theoretisch für sehr flache oberhalb oder unterhalb der Sichtachse befindliche Objekte möglich. Im Falle von mehreren beobachtbaren Referenzecken eines Objekts werden zum SKS-Ursprung näher liegende Punkte als optimaler betrachtet.

Die Entscheidung, ob ein detektierter Eckpunkt \hat{E} auch den entsprechenden realen Eckpunkt E abbildet, wird mit Hilfe der den Sensorsichtbereich beschreibenden Ebenen getroffen (siehe Abschnitt 3.2.3). Liegt ein detektierter Eckpunkt E_d mit einem Mindestabstand innerhalb dieser Ebenen, so wird er als ein wahrer Eckpunkt E betrachtet. Der Mindestabstand bestimmt sich vereinfachend aus der maximalen Halbachsenlänge des F-Ellipsoiden von \hat{E} .

Die Menge der so ermittelten beobachtbaren Objekthecken \mathbb{B} wird anschließend auf Verdeckung durch andere Objekten überprüft. Eine vertikale oder horizontale Kantenverdeckung wird für ein Objekt O_1 angenommen, falls entlang einer Objektkante von O_1 ein oder mehrere näher liegende Objekte unmittelbar angrenzen (siehe Abbildung 5.8(a)). Um die Verdeckung der wahren Eckpunkte besser zu erkennen, werden die Objektausmaße des potentiell verdeckten Objekts, wie in Abbildung 5.8(b) dargestellt, temporär leicht vergrößert. Liegen anschließend beide Eckpunkte einer Kante von O_1 im Verdeckungsbereich von einem oder mehreren Objekten (wie im Beispiel die Punkte \hat{E}_1 und \hat{E}_4), so wird von einer kompletten Kantenverdeckung ausgegangen und die fehlerhaften Kanten-Eckpunkte werden aus \mathbb{B} entfernt. Bei der Verdeckung eines ein-

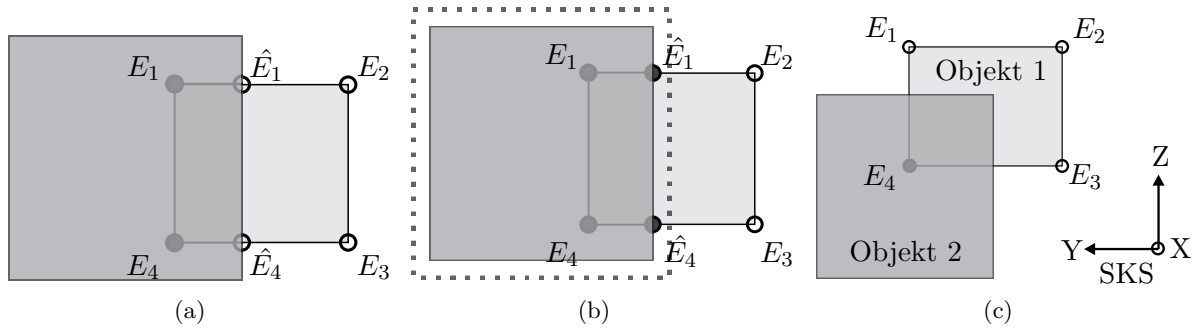


Abbildung 5.8: Verdeckung einer kompletten Objektboxkante (a) und Detektion der Verdeckung durch leichte temporäre Vergrößerung des verdeckenden Objekts (b); partielle Verdeckung einer Objektkante (c)

zelenen Eckpunktes einer Kante, lässt sich der entsprechende Referenzpunkt durch die restlichen Ecken ableiten (Abbildung 5.8(c)).

Da der Rechenaufwand zur Überprüfung der Punktverdeckung durch weitere Objekte mit steigender Objektanzahl exponentiell steigt, wurde eine möglichst effiziente Erkennung implementiert. Für jedes Objekt O werden bei der Erstellung dessen *Verdeckungsraum* \mathcal{R}_V^O und *Objektraum* \mathcal{R}_O^O bestimmt (siehe Abbildung 5.9). Beide Räume lassen sich über eine endliche Menge von sich schneidenden Ebenen beschreiben. Die Ebenen von \mathcal{R}_O bestimmen sich aus den jeweiligen Eckpunkten der sechs Objektseiten. Die Menge der Ebenen $\mathbb{E}_{\mathcal{R}_V}$ zur Definition von \mathcal{R}_V beinhalten jeweils den Koordinatenursprung des SKS sowie zwei Eckpunkte E_x und E_y . Eine Ebene E gehört genau dann zu $\mathbb{E}_{\mathcal{R}_V}$, falls alle Eckpunkte von O auf der gleichen Seite von E liegen bzw. ihre Abstände zu E das gleiche Vorzeichen besitzen.

Liegen die Ebenen in der Hesseschen Normalform vor, so lässt sich für einen beliebigen kartesischen Punkt P über den Normalenvektor bestimmen, ob er innerhalb oder außerhalb von \mathcal{R}_V oder \mathcal{R}_O liegt. Ist ein Punkt P von einem Objekt O verdeckt, so liegt er für alle Ebenen von \mathcal{R}_V^O auf der gleichen Ebenenseite wie der Objektmittelpunkt von O . Um zusätzlich zu überprüfen, ob sich P zwar in \mathcal{R}_V^O aber nicht vor O befindet, muss P für alle Ebenen von \mathcal{R}_O^O auf der entgegengesetzten Ebenenseite zum Koordinatenursprung des SKS liegen.

Nach dem gleichen Verfahren wird auch die Eigenverdeckung von Eckpunkten durch das Objekt selbst überprüft. Aus der so bestimmten Menge \mathbb{B} können die beobachtbaren Eck- und Referenzpunkte abgeleitet werden.

5.3.2 Referenzpunktwechsel

Bei einem Wechsel des optimalen Referenzpunktes im Zuge einer Objektverfolgung, hervorgerufen durch z.B. eine Objektdrehung oder Verdeckung, ist es nötig, den aktuell verfolgten Objektzustand \mathbf{x}' am Referenzpunkt R' , auf den neuen, optimaleren Referenzpunkt R'' nach \mathbf{x}'' zu überführen. Die Kovarianzmatrix \mathbf{P}' ist entsprechend nach \mathbf{P}'' abzubilden (siehe Abbildung 5.10).

Die Koordinaten $(x_{R''}, y_{R''}, z_{R''})$ des neuen Referenzpunktes R'' lassen sich aus dem aktuellen

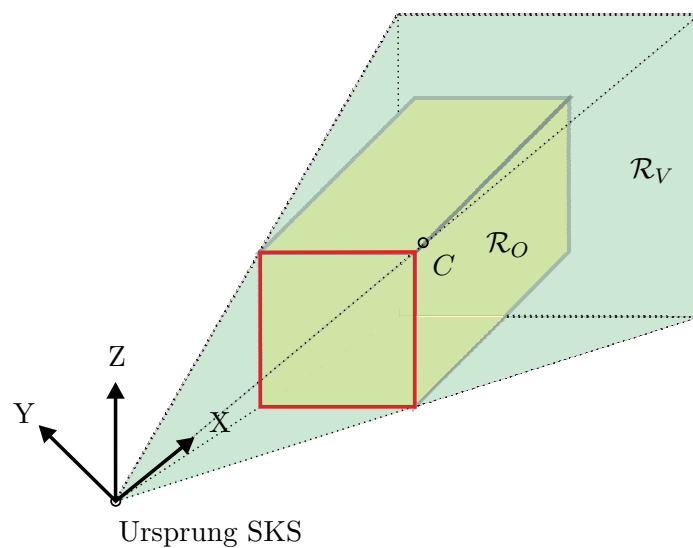


Abbildung 5.9: Verdeckungsraum \mathcal{R}_V (grün) und Objektraum \mathcal{R}_O (gelb) eines Objekts zur Bestimmung von Punktverdeckungen durch das Objekt; \mathcal{R}_V -bestimmende Kanten sind rot dargestellt

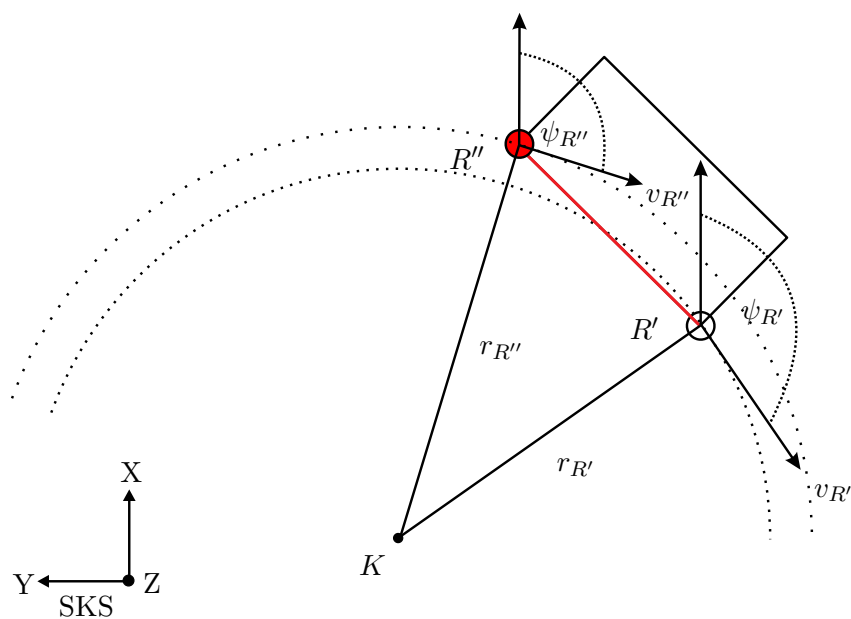


Abbildung 5.10: Wechsel von Referenzpunkt R' nach R'' unter Beachtung der unterschiedlichen Bewegungsrichtungen $\psi_{R'}$ und $\psi_{R''}$ und Abstände $r_{R'}$ und $r_{R''}$ zum Drehmittelpunkt K

5 High Level - Objektverfolgung

Objektzustand \mathbf{X}' , über die vorherige Referenzposition R' , die Objektausmaße (d_x , d_y , d_z) und die Objektausrichtung ψ_O ableiten (Gleichungen 1 und 2 im Anhang).

Da die Beschleunigungen der Bewegungs- und Drehwinkelgeschwindigkeit im Bewegungsmodell des EKF nicht explizit abgebildet sind, wird vereinfachend von einer gleichförmigen Kreisbewegung des Objekts ausgegangen. Dementsprechend können die Bewegungsgeschwindigkeit v und Drehwinkelgeschwindigkeit ω eines Punktes am Objekt über den Kreisradius r zueinander abgebildet werden [88]:

$$v = r\omega . \quad (5.24)$$

Aufgrund der angenommenen gleichförmigen Kreisbewegung besitzen alle Objektpunkte die gleiche Drehwinkelgeschwindigkeit ($\omega_{R'} = \omega_{R''}$). Für eine signifikante Objektdrehgeschwindigkeit ungleich Null mit $|\omega| > \epsilon$ wird vom aktuellen Referenzpunkt R' ausgehend der Mittelpunkt der Kreisbewegung K bestimmt durch

$$K = R' + \frac{v_{R'}}{\omega} \begin{pmatrix} \cos(\psi_{R'}) \\ \sin(\psi_{R'}) \end{pmatrix} \quad \omega > \epsilon . \quad (5.25)$$

Da alle Punkte am Objekt den gleichen Drehmittelpunkt K besitzen, entspricht der Radius $r_{R''}$ dem Kreis für die Kreisbewegung am neuen Referenzpunkt R'' und ergibt sich aus dem Betrag des Differenzvektors zwischen K und R'' . Die Bewegungsgeschwindigkeit $v_{R''}$ und die Bewegungsrichtung $\psi_{R''}$ des Referenzpunktes R'' können so durch

$$v_{R''} = r_{R''}|\omega|, \text{ mit } r_{R''} = |R'' - K| \quad (5.26)$$

und

$$\psi_{R''} = \arctan\left(\frac{y_{R''} - y_K}{x_{R''} - x_K}\right) + \frac{\omega}{|\omega|} \frac{\pi}{2} \quad (5.27)$$

hergeleitet werden.

Aufgrund der analytischen Komplexität der Gleichungen 5.26 und 5.27 wird auf die strenge Bestimmung der Differential-Quotienten zur Ermittlung der Kovarianzmatrix \mathbf{P}'' verzichtet. Stattdessen werden zur Varianzfortpflanzung der Geschwindigkeits- und Drehwinkeltransformation die Differenzen-Quotienten verwendet, da in vielen Fällen

$$\frac{\partial f}{\partial x} \approx \frac{\Delta f}{\Delta x} \quad (5.28)$$

für eine nichtlineare Funktion f und eine messbare Zufallsgröße x gilt [27].

Für einen beliebigen Wert Δx_j , der nahe der jeweiligen Messgenauigkeit angesetzt wird, wer-

den die Differenzen-Quotienten für die jeweiligen Zufallsgrößen x_j bestimmt.

Die Abbildung von P' nach P'' folgt wieder der üblichen Varianzfortpflanzung mit

$$P'' = AP'A^T, \quad (5.29)$$

wobei sich die Überführungsmatrix A aus den Differenzenquotienten der Gleichungen zur Geschwindigkeits- und Drehwinkeltransformation (siehe Gleichungen 5.26 und 5.27) sowie den Differentialquotienten aus den Gleichungen zur Referenzpunktüberführung (siehe Gleichungen 1 und 2 im Anhang) bildet:

$$A = \begin{bmatrix} \frac{\partial x_{R''}}{\partial x} & 0 & 0 & \frac{\partial x_{R''}}{\partial d_x} & \frac{\partial x_{R''}}{\partial d_y} & 0 & \frac{\partial x_{R''}}{\partial \psi} & 0 & 0 \\ 0 & \frac{\partial y_{R''}}{\partial y} & 0 & \frac{\partial y_{R''}}{\partial d_x} & \frac{\partial y_{R''}}{\partial d_y} & 0 & \frac{\partial y_{R''}}{\partial \psi} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{\Delta \psi_{R''}}{\Delta x_{R'}} & \frac{\Delta \psi_{R''}}{\Delta y_{R'}} & 0 & \frac{\Delta \psi_{R''}}{\Delta d_x} & \frac{\Delta \psi_{R''}}{\Delta d_y} & 0 & \frac{\Delta \psi_{R''}}{\Delta \psi} & \frac{\Delta \psi_{R''}}{\Delta v_{R'}} & \frac{\Delta \psi_{R''}}{\Delta \omega} \\ \frac{\Delta v_{R''}}{\Delta x_{R'}} & \frac{\Delta v_{R''}}{\Delta y_{R'}} & 0 & \frac{\Delta v_{R''}}{\Delta d_x} & \frac{\Delta v_{R''}}{\Delta d_y} & 0 & \frac{\Delta v_{R''}}{\Delta \psi} & \frac{\Delta v_{R''}}{\Delta v_{R'}} & \frac{\Delta v_{R''}}{\Delta \omega} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.30)$$

Für den Fall, einer nahezu linearen Objektbewegung mit $|\omega| \leq \varepsilon$ wird für alle Objektpunkte die gleiche Geschwindigkeit und Drehrichtung angenommen:

$$v_{R''} = v_{R'}, \quad \psi_{R''} = \psi_{R'}, \quad \sigma_{v_{R''}}^2 = \sigma_{v_{R'}}^2, \quad \sigma_{\psi_{R''}}^2 = \sigma_{\psi_{R'}}^2 \quad \text{mit } \omega \leq \epsilon. \quad (5.31)$$

Die Zustandsgrößen der Objektdimensionen d_{x_k} , d_{y_k} und d_{z_k} und deren Varianzen bleiben von einem Referenzpunktwechsel unberührt.

5.4 Zusammenfassung

Im vorangehenden Kapitel wurde die Bestimmung der Bewegungsverläufe der Objekthypothesen in Form von Tracks auf der oberen Verarbeitungsebene beschrieben. Für die Zuordnung von Objekthypothesen und Tracks wurde basierend auf einer Assoziationsmatrix ein sensorunabhängiges Assoziationsverfahren entwickelt. Neben den genauigkeitsgewichteten Objektdistanzen können in der Zuordnung auch weitere Objekteigenschaften mit einfließen. Für das Tracking wurde ein erweiterter Kalman Filter zur Modellierung von nicht-linearem Fahrverhalten implementiert. Neben der Position werden zusätzlich die Objektausmaße und die Objektausrichtung mit verfolgt. Einen weiteren Beitrag zur Verbesserung der Tracking-Qualität stellt die Einführung eines Beobachtungsmodells sowie die dynamischen Bestimmung und Nachführung des optimalen Referenzpunktes am Objekt dar.

6 Multilevel-Datenfusion

Die Vorverarbeitung der Einzelsensordaten auf den Prozessierungsebenen wurde in den vorangegangenen Kapiteln dargestellt. Der folgende Abschnitt beschreibt die Verfahren zur Datenfusion zwischen den unterschiedlichen Sensoren auf den drei Verarbeitungsebenen (Punktdatenebene – Low-Level, Objektdatenebene – Mid-Level, Trackebene – High-Level). Durch die Fusion auf der unteren Ebene sollen die Approximationsgenauigkeit der Objektboxausmaße und Ausrichtung erhöht sowie die innerhalb der Messung und Verarbeitung entstehenden Fehler durch die Informationsredundanz der konkurrierenden Sensoren auf den oberen Ebenen korrigiert werden. Durch eine Zusammenführung der Fusionsverfahren der unterschiedlichen Ebenen sollen diese Vorteile kombiniert werden. Wie schon im Kapitel 2.1.1 erwähnt, sind die hier vorgestellten Fusionsverfahren unabhängig von der Sensoranzahl sowie der Funktionsweise der zu integrierenden Sensoren, um die Verfahren und Rückschlüsse auf weitere Sensorsysteme übertragen zu können.

In Abbildung 6.1 sind die Entstehung und Auswirkung der Fehler auf den einzelnen Prozessierungsstufen dargestellt.

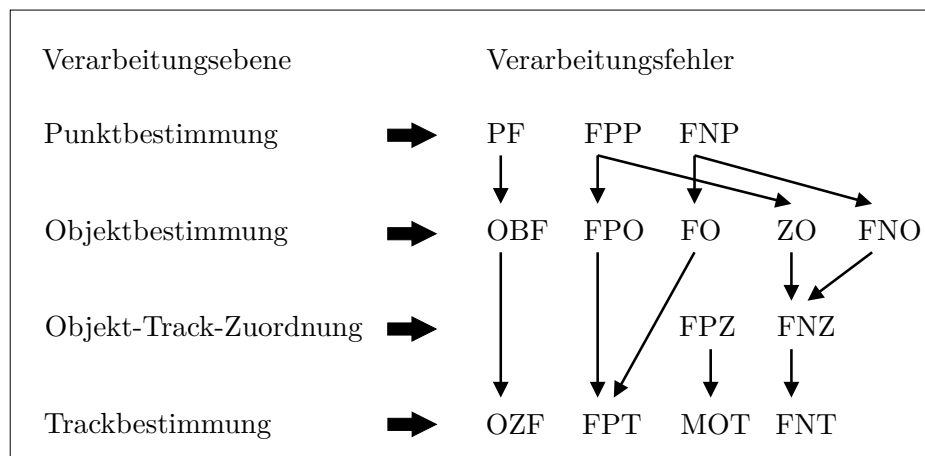


Abbildung 6.1: Entstehung und Fortpflanzung der Detektions- und Prozessierungsfehler von der Punktebene (Punktfehler (PF), Falsch- (FPP) und Fehl-Punkt (FNP)) über die Objektebene (Objektboxfehler (OBF), Falsch- (FPO), Fehl- (FNO), Zusammenfassungs- (ZO) und Fragmentierungs-Objekt (FO)) zur Trackebene (Fehl- (FNZ) und Falschzuordnung (FPZ), Objektzustandsfehler (OZF), Multi-Objekt- (MOT), Fehl- (FNT) und Falsch-Track (FPT))

6.1 Fusion der Punktdaten

Bei der Fusion auf der untersten Verarbeitungsebene (Low-Level-Fusion) werden die gemessenen Punktdaten der Sensoren, so auch im Falle des hier vorgestellten Stereokamera-Laserscanner-Systems, in Form von kartesischen 3D-Koordinaten fusioniert. Da die Punktdaten im Vorfeld schon von ihrem systematischen Fehleranteil befreit und in ein übergeordnetes, gemeinsames Koordinatensystem überführt wurden, ist die Begrifflichkeit der Rohdatenfusion nicht zutreffend und wird somit hier nicht verwendet.

Im Allgemeinen lassen sich die Verfahren der Punktdatenfusion in drei Kategorien unterteilen.

Bei der **punktweisen Fusion** werden die korrespondierenden räumlichen Punktinformationen der verschiedenen Sensoren direkt fusioniert. Als Beispiel sei die Arbeit von [116] genannt, in der die Tiefeninformationen eines kombinierten *Time-of-Flight*-Sensors und einer Stereokamera innerhalb eines globalen Stereo-Matchings zusammengefasst werden. Für den hier vorgestellten Sensoraufbau ist dieser Ansatz jedoch nicht praktikabel, da aufgrund der zu geringen Auflösung der beiden Systeme keine eindeutigen Korrespondenzen zwischen den Objektpunkten bestehen oder sich nur unzureichend genau bestimmen lassen.

Ein weiterer Ansatz zur Fusion der Punktinformationen ist die **gemeinsame Segmentierung**. Dabei wird eine punktweise Fusion vermieden, indem die Punktinformationen der eingehenden Sensoren gemeinsam in die nächsthöhere Prozessierungsstufe, der Segmentierung, eingehen. In [13] werden so die Daten eines Radar-Sensors und eines Laserscanners innerhalb eines räumlichen Belegungsrasters fusioniert. Die gemeinsame Fusion von Stereokamera- und Laserscannerdaten in einem ebenen 2D-Raster wird in [66] und [15] beschrieben.

In der **unterstützenden Segmentierung** werden die Punktinformationen der unterschiedlichen Sensoren nicht wie in der gemeinsamen Segmentierung in einer konkurrierenden Fusion, sondern in einer kooperativen, mehrstufigen Segmentierung kombiniert. So werden in [84] die zuvor bestimmten Laserscanner-Segmente in ein Stereokamera-Tiefenbild abgebildet, um die Objekthypothesen zu verifizieren und zu verfeinern. Ähnliche Ansätze existieren für die Fusion von Laserscanner- [113] oder Radardaten [1] mit einer Monokamera bzw. mit einer Sensor-kombination von Radar und Stereokamera [31], bei dem das Radar die Tiefenbereiche für die Stereodatensegmentierung vorgibt.

Die in dieser Arbeit entwickelte Fusion der Punktdaten entspricht einer gemeinsamen Segmentierung (vorgestellt in [49]). Die Punktinformationen werden hierfür gemeinsam dem Segmentierungsprozess, wie in Abschnitt 4.1 dargestellt, zugeführt (siehe Abbildung 6.2). Grundsätzlich unterscheidet sich hierbei die fusionierende Segmentierung nicht von der Einzelsensorsegmentierung. Die sensorspezifischen Eigenschaften beeinflussen hierbei die fusionierende Segmentierung über das Segmentierungs- und Fehlermodell der Messpunkte. Somit ist auch in nicht überlappenden Detektionsbereichen die fusionierende Segmentierung funktionsfähig. Nach der Fusion werden die fusionierten Segmente entsprechend der beschriebenen Prozesskette (siehe Abbildung 2.3) weiterverarbeitet.

Der Ansatz vermeidet zudem die Diskretisierung der Punktinformationen in planaren Gittern oder räumlichen Rastern (wie in [13, 66, 15] beschrieben), um eine möglichst detailgetreue Objektbeschreibung zu erzielen.

Die Forderung nach einer beliebigen Anzahl an einzubringenden Sensoren sowie der Unabhän-

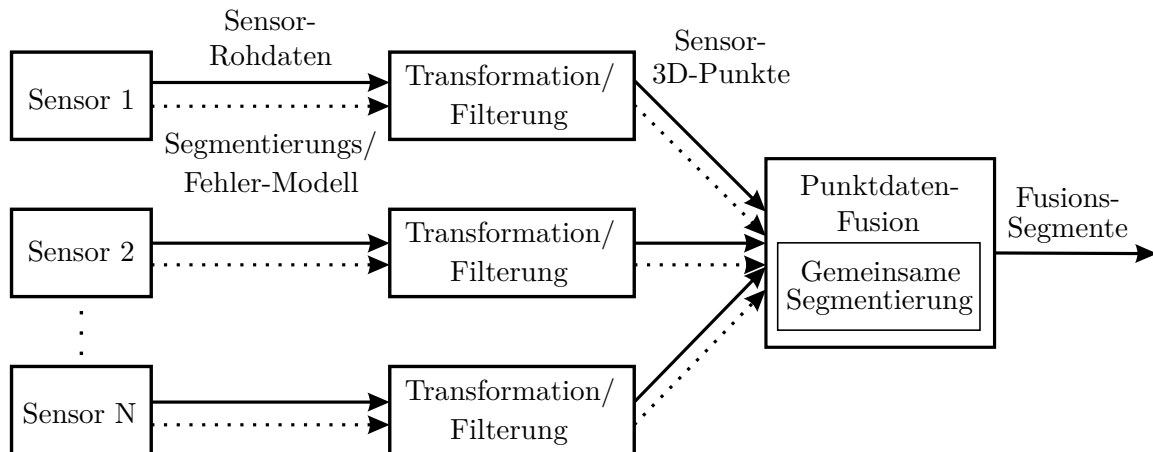


Abbildung 6.2: Low-Level-Fusions-System

gigkeit der Fusion vom Sensortyp wird erfüllt, solange die eingehenden Sensoren umfeldabbildende Messdaten in Form von kartesischen Koordinaten sowie ein Fehler- und Segmentierungsmodell bereitstellen. Weitere Bedingungen für die Low-Level-Fusion stellen die Abbildung der Sensordaten in ein gemeinsames Koordinatensystem (SKS) sowie die zeitlich synchrone Datenakquise durch die Sensoren dar. Andernfalls wäre eine aufwendige punktweise Verfolgung oder Punktzuordnung zwischen aufeinander folgenden Messungen nötig, um eine entsprechende räumliche Interpolation zwischen den korrespondierenden Punktpositionen durchführen zu können.

Eine beispielhafte Darstellung für die Punktdatenfusion des hier vorgestellten Sensorpaares ist in Abbildung 6.3 dargestellt.

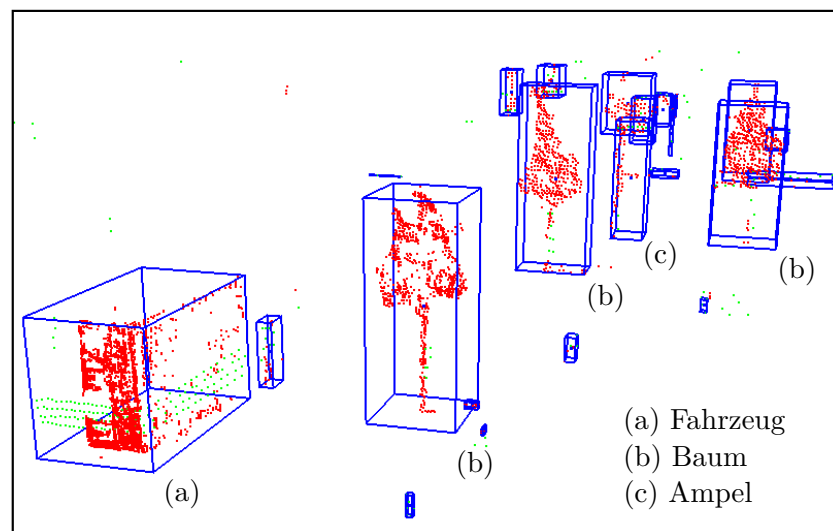


Abbildung 6.3: Fusionierte Segmentierung der roten Stereo- und grünen Laserscannerdaten von der Messszene aus Abbildung 3.4(a)

6.2 Fusion der Objektdaten

Für die Objektdatenfusion, die auch als *Feature-Level-Fusion* oder hier als Mid-Level-Fusion bezeichnet wird, werden die eingehenden Features, hier die Objekthypothesen, bereits im Vorfeld entweder direkt im Sensor oder in einer anschließenden Prozessierungseinheit ermittelt. Als Ergebnis der Fusion ergibt sich ein aus den gewichteten, vorverarbeiteten Eingangsdaten hervorgehendes neues Feature, bzw. eine neue Objekthypothese. Zur Vereinfachung werden im Folgenden Objekthypothesen synonym auch als Objekte bezeichnet.

Im Unterschied zur Low-Level-Fusion wird die Mid-Level-Fusion nach dem Segmentierungsprozess der Einzelsensordaten durchgeführt. Eine im Fahrzeugbereich häufig beschriebene Sensorkonfiguration zur Bestimmung des Fahrumfeldes über eine Mid-Level-Fusion besteht aus einem Laserscanner und einem Mono-Kamerasystem. So werden in [64] und [18] die durch das heterogene Sensorpaar bestimmten Objekthypothesen in einem gemeinsamen Tracking-Filter fusioniert.

In [86] werden die Objektklassifikationen eines Kamera-Laserscanner-Systems als Features über einen Bayes-Klassifikator fusioniert. Die physikalischen Beschreibungen des Objektes werden dabei nur den Laserscannerdaten entnommen.

Die Objektdatenfusion von Laserscanner- und Stereokameradaten wird ebenfalls in [67] beschrieben. Die abgeleiteten Objekthypothesen der Sensoren werden im gemeinsamen Sichtbereich nach einer erfolgreichen paarweisen Zuordnung fusioniert. Hierbei werden die Objekthöhe durch das Stereosystem und die Breiten- und Tiefenangabe durch den Laserscanner bestimmt. Durch den Ausschluss von nur einseitig detektierten Objekten sollen Falschdetektionen vermieden werden.

Eine Auflösung von Fehldetektionen oder der Fragmentierung oder Zusammenfassung von Objekten bieten diese Verfahren jedoch nicht.

Der hier entwickelte Ansatz zur Objektdatenfusion gliedert sich, wie in Abbildung 6.4 dargestellt, in mehrere Schritte. Die detektierten Objekthypothesen der Sensoren werden in einer Clusterung entsprechend der realen Objekte zueinander assoziiert. Anschließend werden die möglichen Segmentierungseffekte, wie das Fragmentieren oder Zusammenfassen oder die Fehl- und Falschdetektion von Objekten, ermittelt und bereinigt. Die assoziierten Objektinformationen werden danach in der eigentlichen Fusion der Zustandsparameter des Objekts zusammengeführt.

Wie schon in der zuvor beschriebenen Low-Level-Fusion müssen die eingehenden Daten der verwendeten Sensoren für die Objektdatenfusion zeitsynchron sein, da die zur zeitlichen Synchronisierung der Objekte nötige Interpolation vermieden wird.

6.2.1 Assoziierung der Multisensor-Objekte

Für die Fusion der Objektinformationen muss im Vorfeld eine klare Zuordnung zwischen den Objekthypothesen der beteiligten Sensoren existieren. Für diese Assoziation wird eine iterative Clusterung der Objektinformationen vorgenommen. Ein *Cluster* besitzt dabei selbst eine Objektstruktur, die sich aus der Menge der aktuell zum Cluster assoziierten Sensorobjekte ableitet. Zusätzlich enthält jeder Cluster in seiner *Sensorkonfiguration* die eindeutige Zuordnung zwischen den einzelnen Sensoren und ihren zugehörigen Cluster-Objekten. Kann ein Sensor kein Objekt

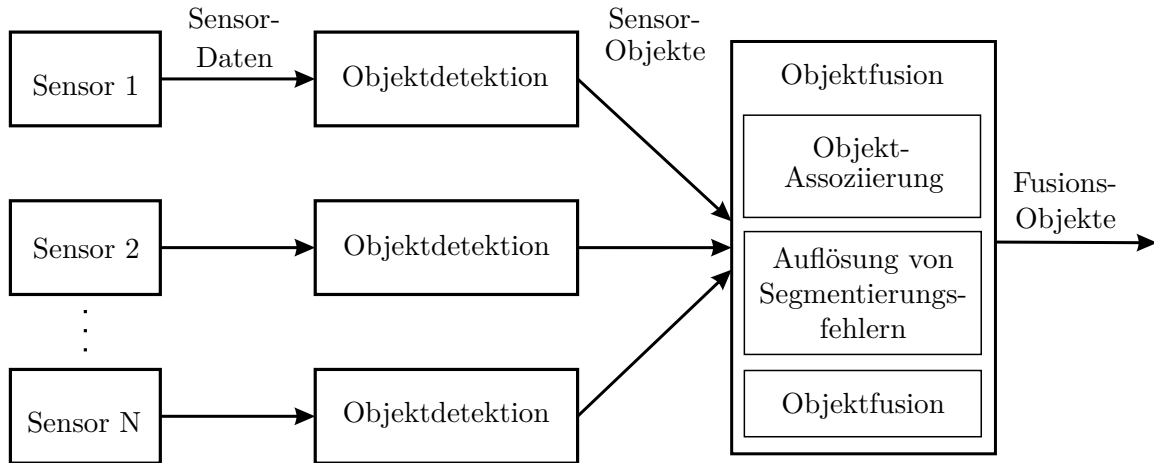


Abbildung 6.4: Mid-Level-Fusions-System

für einen Cluster beisteuern, so wird diese Information ebenfalls in der Sensorkonfiguration hinterlegt.

Die Assoziierung wird analog dem im Abschnitt 5.1 beschriebenen Verfahren über eine Assoziationsmatrix durchgeführt. Die Spaltenelemente repräsentieren die einzelnen Cluster und die Zeilenelemente die zu assoziierenden Objekte der Sensoren. Die für die Aufstellung der Assoziationsmatrix G nötige Distanzfunktion zur Bestimmung des Abstandes d_{ij} zwischen einem Cluster i und einer Objekthypothese j eines Sensors entspricht der in Gleichung 5.10 definierten Funktion. Der Korrelationsterm aus Gleichung 5.1 entfällt jedoch, da die sensorspezifischen Ähnlichkeitsmaße sich nicht auf alle Sensortypen übertragen lassen. Matrix G bestimmt sich demnach mit:

$$G_{ij} = e^{-d_{ij}/2\lambda^2} . \quad (6.1)$$

Zu Beginn wird die Clusterliste mit der Objektliste eines willkürlichen Sensors initialisiert, wobei jede Objekthypothese einen eigenen Cluster initiiert. Sukzessiv werden die Objektmengen der eingehenden Sensoren mit der aktuellen Clusterkonfiguration abgeglichen. Bei einer Erweiterung oder Reduzierung des Clusters um eine Objekthypothese werden die Clusterparameter neu bestimmt.

Die Ermittlung der Clusterparameter für mehrere Sensorobjekte entspricht hierbei einer reduzierten Zustandsfusion entsprechend der Gleichung 6.19. Dabei werden die vier Referenzpunkte und der Objektmittelpunkt sowie die Objektdimensionen und die zugehörigen Kovarianzen der clusterbildenden Objekte der Sensoren zusammengefasst. Befindet sich nach einem wiederholten Clusterungsdurchlauf ein zu assoziierendes Objekt bereits im entsprechenden Cluster, so wird der Cluster temporär um das entsprechende Objekt reduziert, um die Distanzbestimmung nicht zu verzerren. Cluster, die nach einer Assoziation keine Objekte mehr enthalten, werden aus der Clusterliste entfernt. Eine Überprüfung auf Verdeckung wird nicht vorgenommen, da durch die

relativ ähnlichen Perspektiven der Fahrzeugsensoren eine Objektverdeckung für alle Sensoren gleichermaßen eintreten würde.

Die Clusterung wird solange für alle Sensoren wiederholt, bis sich die gesamte Clusterliste über einen kompletten Clusterungsdurchlauf über alle Sensoren nicht mehr verändert.

Nach dem Abschluss der Clusterung werden die aufgrund fehlender Objektzuordnungen unvollständigen Cluster anhand ihrer Sensorkonfiguration untersucht. Kann ein Sensor kein Objekt zum Cluster zuordnen, wird die Beobachtbarkeit des Clusterobjekts für den Sensor überprüft (siehe Abschnitt 3.2.3) und entsprechend in der Sensorkonfiguration des Clusters hinterlegt.

6.2.2 Auflösung von Fragmentierung und Zusammenfassung

Durch die im Abschnitt 4.2.4 beschriebenen Fehler in der Segmentierung bzw. Detektion kann es beispielsweise zu einer Fragmentierung eines Objekts in mehrere Objekthypothesen (FO) oder zur Zusammenfassung mehrerer Objekte zu einer Objekthypothese (ZO) kommen. Diese Effekte können durch die Informationsredundanz mehrerer, das gleiche Objekt beobachtender Sensoren erkannt und kompensiert werden.

Die Grundlage zur Erkennung dieser Segmentierungseffekte bildet die in der Clusterung bestimmte Sensorkonfiguration eines Clusters. Als Indikator wird eine widersprüchliche Aussage zwischen den Sensoren innerhalb einer Sensorkonfiguration eines Clusters gewählt. Ein Konflikt-Cluster liegt vor, wenn die Menge der beteiligten Sensoren, die in der Lage sind das Clusterobjekt zu beobachten, sich in die Untermengen \mathbb{D} und $\bar{\mathbb{D}}$ der detektierenden und nicht-detektierenden Sensoren aufteilt. Wird dieser Widerspruch für einen Cluster C_1 durch einen FO- oder ZO-Effekt herbeigeführt, so sollte zusätzlich für mindestens einen Sensor S_1 aus $\bar{\mathbb{D}}$ eine zu einem zweiten Cluster C_2 zugeordnete Objektdetektion existieren, die den Cluster C_1 überlappt (siehe Abbildungen 6.6 und 6.7). Eine Überlappung eines Clusters C_1 durch ein Sensorobjekt eines Clusters C_2 wird als wahrscheinlich angenommen, falls der Mittelpunkt von C_2 innerhalb des Sensorobjektes liegt (Bestimmung über den Objektraum \mathcal{R}_O siehe Abschnitt 5.3.1).

Um zu entscheiden, ob in diesem Fall die beiden Cluster C_1 und C_2 ein und dasselbe Objekt abbilden (FO) oder ob ein oder mehrere Sensoren zwei reale Objekte, die durch die Cluster C_1 und C_2 repräsentiert werden, zusammengefasst haben (ZO), werden die einzelnen konkurrierenden Aussagen der Sensoren gegeneinander abgewogen.

Alle Sensoren, die ein überlappendes Objekt über C_1 und C_2 detektiert haben, stimmen für den FO-Effekt und bilden die Menge \mathbb{U} mit

$$\mathbb{U} \subseteq (\mathbb{D}_{C_1} \cap \bar{\mathbb{D}}_{C_2}) \cup (\bar{\mathbb{D}}_{C_1} \cap \mathbb{D}_{C_2}) . \quad (6.2)$$

Die Sensoren, die sowohl für C_1 als auch für C_2 ein Objekt zugeordnet haben, werden für den ZO-Effekt und bilden die Menge $\bar{\mathbb{U}}$:

$$\bar{\mathbb{U}} \subseteq \mathbb{D}_{C_1} \cap \mathbb{D}_{C_2} . \quad (6.3)$$

Für jeden Sensor X aus \mathbb{U} und $\bar{\mathbb{U}}$ lässt sich eine Wahrscheinlichkeit $P_X(F)$ über die binäre

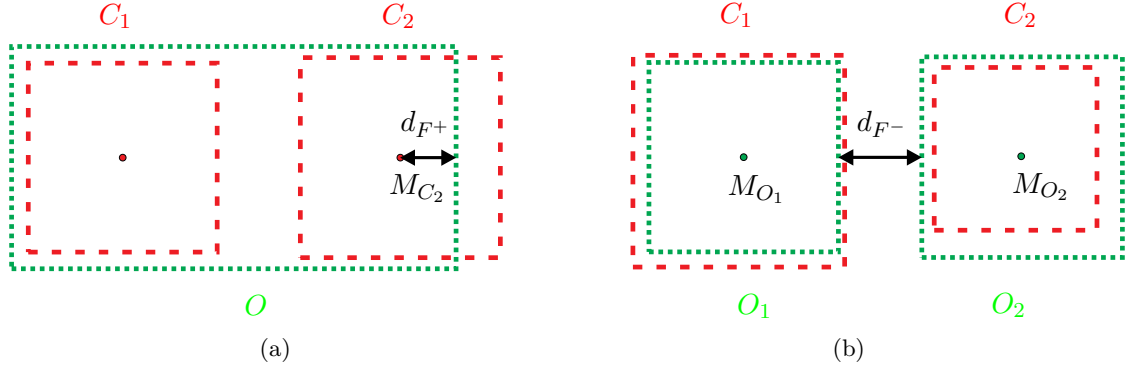


Abbildung 6.5: Bestimmung der Überlappungsdistanz d_{F+} zwischen dem Objekt O und dem Cluster C_2 (a); Bestimmung der Trennungsdistanz d_{F-} zwischen den Objekten O_1 und O_2 (b)

Zufallsgröße F (Objekt ist fragmentiert F_+ und Objekt ist nicht fragmentiert F_-) bestimmen. Die Grundlage hierfür liefert der Grad der Überlappung oder Trennung der Sensorobjekte aus \mathbb{U} und $\bar{\mathbb{U}}$ über C_1 und C_2 .

Ein Sensor X aus \mathbb{U} spricht aufgrund einer detektierten Überlappung zweier Cluster für eine Objektfragmentierung (F_+) in den Konflikt-Clustern. Hierfür wird analog zur Segmentierungswahrscheinlichkeit (SW) (siehe Abschnitt 4.1.3) der Vektor \mathbf{D}_{F+} des kürzesten Abstands d_{F+} zwischen dem überlappten Clustermittelpunkt M_{C_2} und der Objektbox des Überlappungsobjekts O von Sensor X genutzt. Über die Verteilungsfunktion der Standardnormalverteilung $\Phi(z)$ und der Punktgenauigkeit σ_{M_C} von M_{C_2} in Richtung \mathbf{D}_{F+} bis zur Objektgrenze von O kann die Wahrscheinlichkeit der Überlappung mit

$$P^-(F_+) = \begin{cases} \Phi\left(\frac{d_{F+}}{\sigma_{M_C}}\right) & , \text{ mit } d_{F+} \geq 0 \\ 1 - \Phi\left(-\frac{d_{F+}}{\sigma_{M_C}}\right) & , \text{ mit } d_{F+} < 0 \end{cases} \quad (6.4)$$

und somit der Fragmentierung abgeschätzt werden. Je größer die Überlappungsdistanz d_{F+} , desto höher ist somit auch die Überlappungswahrscheinlichkeit $P^-(F_+)$ (siehe Abbildung 6.5(a)). Die Überlappungsdistanz d_{F+} wird vom Betrag negativ, sofern der Clustermittelpunkt M_{C_2} nicht mehr innerhalb des Objekts O liegt.

Neben der Überlappungswahrscheinlichkeit $P^-(F_+)$ eines Objektes O über einen Clustermittelpunkt M_C wird zusätzlich das Segmentierungsverhalten an der Stelle des Mittelpunkts M_O von O bei der Bestimmung der Fragmentierungswahrscheinlichkeit $P(F_+)$ mit einbezogen. Liegt die SW $P(S_{M_O}^X)$ zwischen 75 % und 50 % so wird eine erhöhte Fragmentierungswahrscheinlichkeit angenommen. Für eine SW über 75 % erhöht sich die Wahrscheinlichkeit einer Objektzusammenfassung:

$$P(F_+) = (P^-(F_+) - \frac{1}{2})(\frac{3}{2} - P(S_{M_O}^X)) + \frac{1}{2}. \quad (6.5)$$

Analog lässt sich für jeden Sensor Y aus $\bar{\mathbb{U}}$, der zwei Objekte O_1 und O_2 detektiert hat, die Nicht-Fragmentierungswahrscheinlichkeit $P(F_-)$ ableiten. Als Maß für eine korrekt erkannte Trennung zweier Objekte O_1 und O_2 wird der minimale Objektabstand als Vektor \mathbf{D}_{F_-} und dessen Betrag d_{F_-} (Trennungsdistanz) zwischen O_1 und O_2 herangezogen (siehe Abbildung 6.5(b)). Die Streuung σ_{M_O} wird gleich der höheren Streuung der beiden Objektmittelpunkte M_{O_1} und M_{O_2} gesetzt. Über die Verteilungsfunktion der Standardnormalverteilung $\Phi(z)$ und der Punktgenaugigkeit σ_{M_O} in Richtung des Trennungsvektors \mathbf{D}_{F_-} kann die Trennungswahrscheinlichkeit $P^-(F_-)$ mit

$$P^-(F_-) = \begin{cases} \Phi\left(\frac{d_{F_-}}{\sigma_{M_O}}\right) & , \text{ mit } d_{F_-} \geq 0 \\ 1 - \Phi\left(-\frac{d_{F_-}}{\sigma_{M_O}}\right) & , \text{ mit } d_{F_-} < 0 \end{cases} \quad (6.6)$$

bestimmt werden. Im Falle einer Überlappung von O_1 und O_2 wird der Betrag von d_{F_-} negativ gesetzt. Größere Abstände erhöhen und kleiner Abstände verringern somit die Trennungswahrscheinlichkeit. Für $d_{F_+} = d_{F_-} = 0$ und gilt $P^-(F_+) = P^-(F_-) = 0.5$.

Die resultierende Trennungswahrscheinlichkeit $P^-(F_-)$ wird aus den zuvor genannten Gründen ebenfalls mit der SW $P(S_{M_{O_1O_2}}^Y)$ am Mittelpunkt

$$M_{O_1O_2} = \frac{M_{O_1} + M_{O_2}}{2} \quad (6.7)$$

zwischen den Mittelpunkten M_{O_1} und M_{O_2} von O_1 und O_2 gewichtet:

$$P(F_-) = (P^-(F_-) - \frac{1}{2})(P(S_{M_{O_1O_2}}^Y)) + \frac{1}{2}. \quad (6.8)$$

Die komplementären Wahrscheinlichkeiten $P(F_+)$ und $P(F_-)$ lassen sich aufgrund ihrer Binarität für jeden der M Sensoren aus \mathbb{U} und $\bar{\mathbb{U}}$ bestimmen und über eine Multiplikation zusammenfassen und vergleichen:

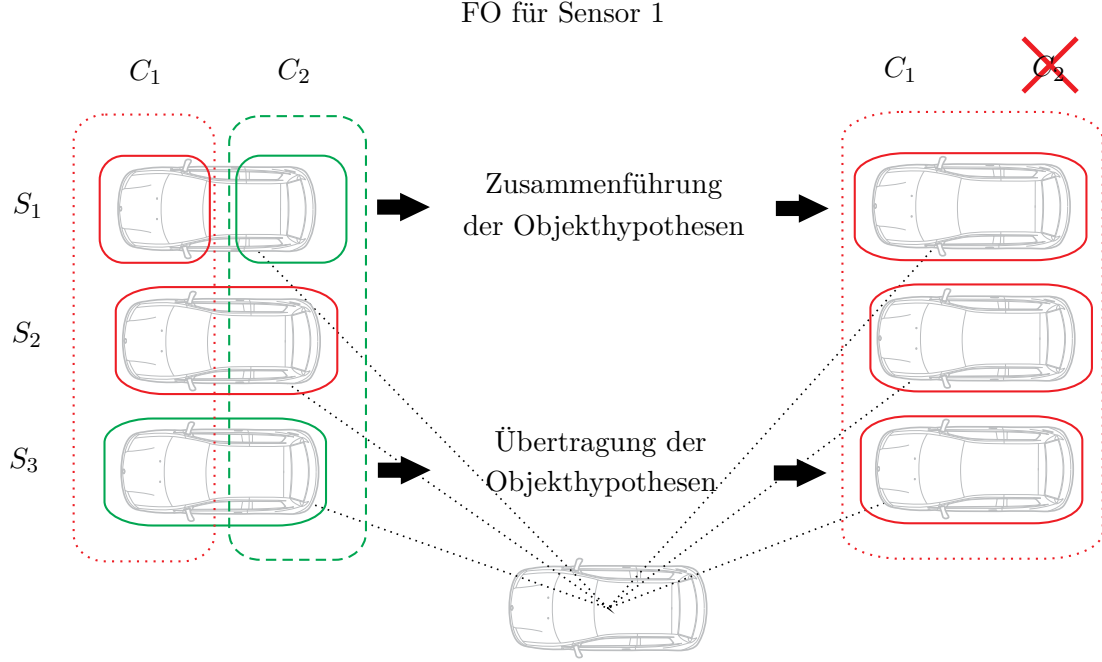


Abbildung 6.6: Auflösung der FO für drei Sensoren; Objektfragmentierung in der Objekterkennung bei Sensor 1 innerhalb der Cluster 1 und 2 (linke Seite); Auflösung der Fragmentierung durch Zusammenführung der Objekthypothesen in einem Cluster (rechte Seite)

$$\begin{aligned}
 \prod_{i=1}^M P_{X_i}(F_+) &> \prod_{i=1}^M P_{X_i}(F_-), \text{ mit } X^i \in (\mathbb{U} \cup \bar{\mathbb{U}}) \rightarrow \text{FO} \\
 \prod_{i=1}^M P_{X_i}(F_+) &< \prod_{i=1}^M P_{X_i}(F_-), \text{ mit } X^i \in (\mathbb{U} \cup \bar{\mathbb{U}}) \rightarrow \text{ZO} \\
 \prod_{i=1}^M P_{X_i}(F_+) &= \prod_{i=1}^M P_{X_i}(F_-), \text{ mit } X^i \in (\mathbb{U} \cup \bar{\mathbb{U}}) \rightarrow \text{Keine Entscheidung.} \quad (6.9)
 \end{aligned}$$

Im Falle einer Objektfragmentierung werden die Objektfragmente eines Sensors wieder fusioniert, die restlichen Sensorobjekte beider Cluster in einem Cluster vereint und der überflüssige, leere Cluster entfernt (siehe Abbildung 6.6).

Zur Auflösung einer erkannten Objektzusammenfassung bestehen zwei Handlungsoptionen. Entweder werden die zusammengefassten Objekte aus dem Cluster entfernt oder entsprechend der angenommenen Objektverteilung zerteilt. Da sich die Genauigkeiten für die Parameter der aus der Teilung resultierenden Objekte schwer abschätzen lassen, besteht die Gefahr einer falschen Gewichtung in der anschließenden Zustandsfusion (siehe Abschnitt 6.2.4). Dementspre-

ZO für Sensor 2

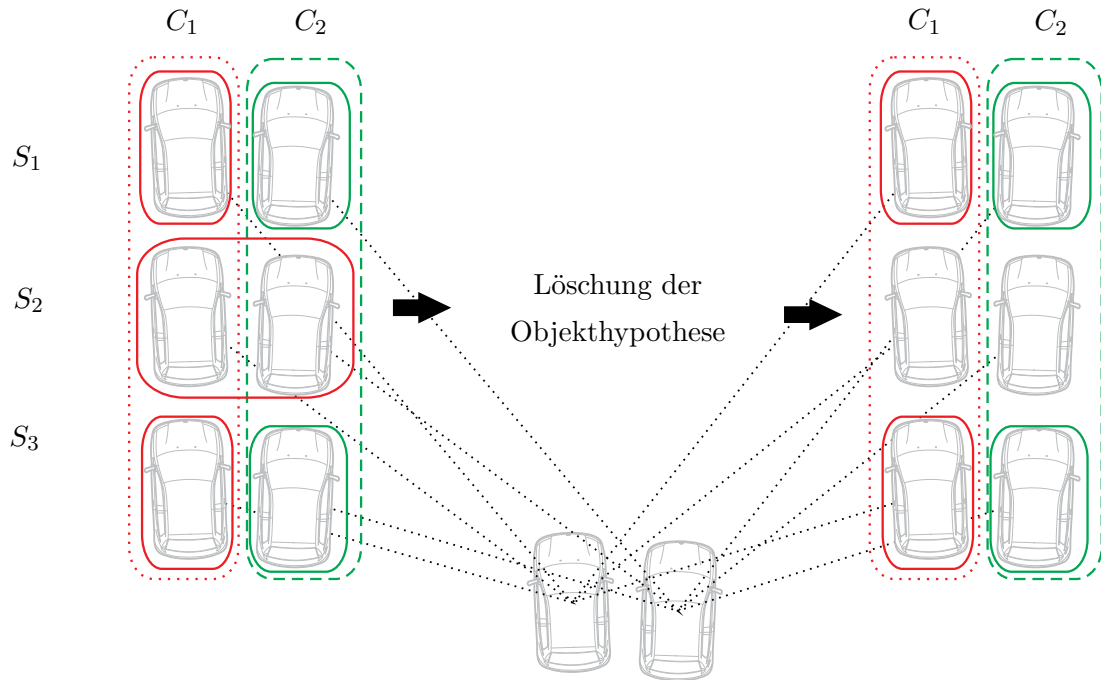


Abbildung 6.7: Auflösung der ZO für drei Sensoren; Objektzusammenfassung in der Objekterkennung bei Sensor 2 innerhalb der Cluster 1 und 2 (linke Seite); Auflösung der Zusammenfassung durch Entfernung der Objekthypothese aus den Clustern (rechte Seite)

chend konservativ wird das zusammengefasste Objekt gänzlich aus dem Cluster entfernt (siehe Abbildung 6.7).

Die FO/ZO-Erkennung und Auflösung wird für alle Cluster der Clusterliste sukzessiv durchgeführt und solange wiederholt, bis keine Veränderung in der Clusterliste mehr auftritt. Durch die wiederholte Ausführung können auch mehrfach fragmentierte oder zusammengesetzte Objekte erkannt werden.

6.2.3 Auflösung von Fehl- und Falschdetektionen

Nach der FO/ZO-Erkennung werden die restlichen Cluster mit unaufgelösten widersprüchlichen Sensorkonfigurationen überprüft, ob die Konflikte durch eine Fehl- oder Falschdetektion verursacht wurden (siehe Abschnitt 4.2.4).

In [67] werden zur Vermeidung von Falschdetektionen nur Objekthypothesen zur Fusion zugelassen, die von beiden beteiligten Sensoren durch eine Hypothesenzuordnung unterstützt werden, wodurch jedoch Fehldetektionen nicht erkannt werden. Die Existenzwahrscheinlichkeit eines Objektes wird in [79] mittels eines auf der Dempster Shafer Theorie [99] basierten Existenztrackers

über die Zeit verfolgt (JIPDA-DS).

Da bei der Prozessierung in der Objektebene bewusst auf den Zugriff auf historische Informationen verzichtet wird, um sich selbst verstärkende Fehler zu vermeiden, können nur die aktuellen Sensordaten zur Existenzvalidierung eines Konflikt-Clusters herangezogen werden. Um für einen Konfliktcluster C zu entscheiden, ob er falsch detektiert wurde (FPO) oder ob einige Sensoren für den Cluster kein Objekt detektieren bzw. zuordnen konnten (FNO), werden die beiden komplementären Verbundwahrscheinlichkeiten $P(E_+|D_+)$ und $P(E_-|D_+)$ für C bestimmt und gegenübergestellt.

Die Wahrscheinlichkeit $P(E_+|D_+)$, dass falls eine Objekthypothese detektiert wurde (D_+), dieses Objekt auch existiert (E_+) sowie die Wahrscheinlichkeit $P(E_-|D_+)$, dass falls eine Objekthypothese detektiert wurde, dieses Objekt nicht existiert (E_-), können aus den Informationen der Konfliktsensoren der Mengen \mathbb{D} und $\bar{\mathbb{D}}$ abgeleitet werden. Die beiden diskreten Zufallsgrößen E und D sind hierbei binär.

Da sich diese Wahrscheinlichkeiten nicht direkt bestimmen lassen, wird das Bayes Theorem (z.B. siehe [9]) angewendet, um diese aus den bestimmmbaren Wahrscheinlichkeiten mit

$$P(E_+|D_+) = \frac{P(D_+|E_+)P(E_+)}{P(D_+|E_+)P(E_+) + P(D_+|E_-)P(E_-)} \quad (6.10)$$

und

$$P(E_-|D_+) = \frac{P(D_+|E_-)P(E_-)}{P(D_+|E_+)P(E_+) + P(D_+|E_-)P(E_-)} \quad (6.11)$$

abzuleiten, wobei gilt:

$$P(E_-|D_+) = 1 - P(E_+|D_+) . \quad (6.12)$$

Zur Bestimmung von $P(E_+|D_+)$ (Posteriori-Wahrscheinlichkeit) werden entsprechend Gleichung 6.10 die Wahrscheinlichkeit $P(D_+|E_+)$ (Likelihood) für eine positive Detektion, falls das Objekt existiert und die Priori-Wahrscheinlichkeit $P(E_+)$ für das allgemeine Auftreten von Objekten benötigt. Die Priori-Wahrscheinlichkeit wird dabei aus Mangel an Informationen auf einen konstanten neutralen Wert von 0.5 gesetzt. Es bietet sich jedoch an, Infrastrukturinformationen als Priori-Information zu nutzen. So könnten befahrbare Flächen bzw. die eigene Route mit einer höheren Wahrscheinlichkeit für eine Objektexistenz bewertet werden, um somit z.B. die Wahrscheinlichkeit für eine Fehldetektion auf einer unbefahrten Fläche stärker zu gewichten.

Für die Likelihood $P(D_+|E_+)$ eines Sensors X aus der Menge \mathbb{D} eines Clusters C wird die jeweilige Segmentierungswahrscheinlichkeit $P(S_{M_C}^X)$ dieses Sensors (siehe Abschnitt 4.1.3) an der Stelle des Mittelpunktes M_C von C gewählt. Für einen Sensor aus der Menge $\bar{\mathbb{D}}$ entspricht hingegen die jeweilige Segmentierungswahrscheinlichkeit der Likelihood $P(D_+|E_-)$, dass der Sensor positiv detektiert, sofern kein Objekt an der Position existiert.

Aufgrund der Binarität der Zufallsgrößen E und D gilt zudem

$$P(D_+|E_-) = 1 - P(D_+|E_+) , \quad (6.13)$$

wodurch für alle N Sensoren aus \mathbb{D} und $\bar{\mathbb{D}}$ jeweils die Likelihoods $P(D_+|E_+)$ und $P(D_+|E_-)$ berechenbar sind.

Um nun die jeweiligen Likelihood-Informationen der Sensoren in den gemeinsamen Posteriori-Wahrscheinlichkeiten $P(E_+|D_+^1, D_+^2, \dots, D_+^N)$ und $P(E_-|D_+^1, D_+^2, \dots, D_+^N)$ über die Bayes-Regel zu bündeln, stehen prinzipiell zwei Verfahren zur Auswahl. Im in [73] beschriebenen *Independent Opinion Pool* (IOP) werden die zuvor bestimmten einzelnen Posteriori-Wahrscheinlichkeiten der Sensoren über ein Produkt kombiniert. Hierbei wird von einer Unabhängigkeit der Priori-Informationen ausgegangen. Ist diese Unabhängigkeit wie im vorliegenden Fall nicht gegeben, so kann es zu einer Verzerrung der abgeleiteten globalen Posteriori-Wahrscheinlichkeit kommen. In diesem Fall sollten, wie im *Independent Likelihood Pool* (ILP) beschrieben [73], die Likelihoods der Sensoren separat mit

$$\begin{aligned} P(D_+^1, D_+^2, \dots, D_+^N|E_+) &= P(D_+^1|E_+)P(D_+^2|E_+), \dots, P(D_+^N|E_+) \\ P(D_+^1, D_+^2, \dots, D_+^N|E_-) &= P(D_+^1|E_-)P(D_+^2|E_-), \dots, P(D_+^N|E_-) \end{aligned} \quad (6.14)$$

fusioniert werden, um sie anschließend mit der gemeinsamen Priori-Information zu bewerten und unter Einbeziehung der Regel von Bayes nach den gesuchten Posteriori-Wahrscheinlichkeiten abzuleiten:

$$\begin{aligned} P(E_+|D_+^1, D_+^2, \dots, D_+^N) &= \frac{P(D_+^1, D_+^2, \dots, D_+^N|E_+)P(E_+)}{P(D_+^1, D_+^2, \dots, D_+^N)} \\ P(E_-|D_+^1, D_+^2, \dots, D_+^N) &= \frac{P(D_+^1, D_+^2, \dots, D_+^N|E_-)P(E_-)}{P(D_+^1, D_+^2, \dots, D_+^N)} . \end{aligned} \quad (6.15)$$

Da der Normierungsterm $P(D_+^1, D_+^2, \dots, D_+^N)$ für die beiden fusionierten Posteriori-Wahrscheinlichkeiten identisch ist, kann auf eine explizite Berechnung verzichtet werden und die höhere der beiden Gesamt-Posteriori-Wahrscheinlichkeiten entscheidet über eine Fehldetektion oder eine Falschdetektion mit

$$\begin{aligned} P(E_+|D_+^1, D_+^2, \dots, D_+^N) &> P(E_-|D_+^1, D_+^2, \dots, D_+^N) \rightarrow \text{FNO} \\ P(E_+|D_+^1, D_+^2, \dots, D_+^N) &< P(E_-|D_+^1, D_+^2, \dots, D_+^N) \rightarrow \text{FPO} \\ P(E_+|D_+^1, D_+^2, \dots, D_+^N) &= P(E_-|D_+^1, D_+^2, \dots, D_+^N) \rightarrow \text{Keine Entscheidung.} \end{aligned} \quad (6.16)$$

In [87] wird ebenfalls angemerkt, dass sich in den meisten Fällen der Multisensor-Fusion eine Bündelung der Sensorinformationen über ILP besser eignet, da sich die Priori-Informationen in der Regel vom gleichen Ursprung ableiten und als nicht unabhängig betrachtet werden können.

Sensor	$P(D_+ E_+)$	$P(D_+ E_-)$
X	0.7	0.3
Y	0.8	0.2
Z	0.2	0.8

Tabelle 6.1: Beispielhafte Verteilung der Likelihoods für einen Objektcluster der Sensoren X , Y und Z

$P(E_+)$	$P(E_+ D_+^1, D_+^2, D_+^3) \rightarrow \text{FNO}$	$P(E_- D_+^1, D_+^2, D_+^3) \rightarrow \text{FPO}$
0.5	0.7	0.3
0.2	0.36	0.64
0.7	0.97	0.03

Tabelle 6.2: Die über ILP ergebenden Posteriori-Wahrscheinlichkeiten aus den Sensor-Likelihoods von Tabelle 6.1 für verschiedene Priori-Wahrscheinlichkeiten

Bei einem erkannten FNO-Effekt wird der Konflikt-Cluster C als korrekt detektiert angenommen. Der Konflikt entstammt dementsprechend einer fehlenden Objektdetektion eines Sensors im Cluster. Somit bedarf der Konflikt-Cluster C keiner Änderung. Wird von einer fehlerhaften Objektdetektion (FPO) ausgegangen, so wird der entsprechende Konflikt-Cluster aus der Clusterliste entfernt.

Beispiel

Die drei Sensoren X , Y und Z besitzen einen gemeinsamen Detektionsbereich und sind wie beschrieben in der Lage Objekte zu detektieren. Für einen Konfliktcluster C treten dabei die in Tabelle 6.1 dargestellten Likelihoods auf.

Tabelle 6.2 zeigt die Auswirkungen der gewählten Priori-Wahrscheinlichkeiten auf die über eine Fusion mit ILP ergebenden Posteriori-Wahrscheinlichkeiten auf. Die von der neutralen Gewichtung von $P(E_+) = P(E_-) = 0.5$ abweichenden Priori-Informationen bewirken entgegengesetzte Ergebnisse in der Konfliktauflösung. Genaue Priori-Informationen, wie z.B. Karteninformationen, können demnach das Fusionsergebnis entscheidend beeinflussen und Falsch- und Fehldetektionen reduzieren.

6.2.4 Zustandsfusion

Ähnlich der Punktdatenfusion kann die Objektdatenfusion der Objektparameter in Form der Messwertvektoren und ihrer Varianzen in zwei verschiedene Formen unterteilt werden.

In der **objektweisen Fusion** werden die Messwertvektoren $z_1 - z_N$ der Objekthypothesen von N Sensoren sowie die Messgenauigkeiten $R_1 - R_N$ über eine Gewichtung direkt fusioniert.

Eine nach der Methode der kleinsten Quadrate bestimmte Optimierung der Fusionsgewichtung (WLS – *weighted least squares*) finden sich beispielsweise in [43] und [91].

In [45] werden ebenfalls zwei auf WLS basierte Verfahren für eine Objektdatenfusion vorgeschlagen, die zudem unterschiedliche Messwertvektoren der Sensoren und Korrelationen im Messrauschen der verschiedenen Sensoren mit berücksichtigen. Nachteilig wirkt sich jedoch hier der hohe Rechenaufwand durch die Invertierung von hochdimensionalen Matrizen aus. Sowohl die in [43] als auch die in [45] vorgestellten Verfahren sind funktional äquivalent zur zentralisierten Kalman-Fusion.

Im **gemeinsamen Tracking** werden die einzelnen Objektinformationen in der nächsthöheren Prozessierungsstufe, dem Tracking, zusammengeführt. Liegt ein asynchroner Sensoraufbau vor und werden die Daten zu verschiedenen Zeitpunkten akquiriert, so können diese beispielsweise in einem asynchronen Kalman-Filter sukzessiv fusioniert werden. Bei einem zeitsynchronen Sensorsystem wird dabei häufig die inverse Form des Kalman-Filter, der Informationsfilter (siehe [75, 17]), genutzt, um die Informationen der beteiligten Sensoren gemeinsam in die Zustandsschätzung des globalen Tracks (Systemtrack) mit einzubeziehen [18].

Für diese Arbeit wird der Ansatz aus [43] verwendet (*Measurement Fusion Method 2*), da wie hierfür gefordert und bedingt durch das Fusions-Framework nur typgleiche Objektdaten mit einheitlichen Messwertvektoren fusioniert werden. Der Ansatz minimiert hierbei den mittleren quadratischen Fehler und ist funktional äquivalent zur Fusion in einem gemeinsamen Beobachtungsvektor eines Kalman Filters. Analog zur Clusterbildung (siehe Abschnitt 6.2.1) setzt sich der Messwertvektor \mathbf{Z} aus den sechs Parametern des Objektmittelpunkts (x, y, z) und der Objektausmaße (d_x, d_y, d_z) zusammen. Zusätzlich wird \mathbf{Z} um die Objektausrichtung κ erweitert:

$$\mathbf{Z} = \begin{bmatrix} x & y & z & d_x & d_y & d_z & \kappa \end{bmatrix}^T. \quad (6.17)$$

Der fusionierte Messwertvektor \mathbf{Z} und die Messvarianzmatrix \mathbf{R} für N eingehende Sensoren zum Zeitpunkt k ergeben sich demnach zu:

$$\mathbf{R}_k = \left[\sum_{i=1}^N \mathbf{R}_{i_k}^{-1} \right]^{-1} \quad (6.18)$$

$$\mathbf{Z}_k = \left[\sum_{i=1}^N \mathbf{R}_{i_k}^{-1} \right]^{-1} \sum_{i=1}^N \mathbf{R}_{i_k}^{-1} \mathbf{Z}_{i_k}. \quad (6.19)$$

Der Fusionsansatz ist somit wie gefordert unabhängig von der Anzahl der eingebundenen Sensorquellen und des Sensortyps. In Abbildung 6.8 sind für das Fahrzeug aus der Messszene aus Abbildung 3.4(a) die Ergebnisse der Einzelsegmentierungen der Sensoren sowie die daraus resultierende Objektdatenfusion dargestellt.

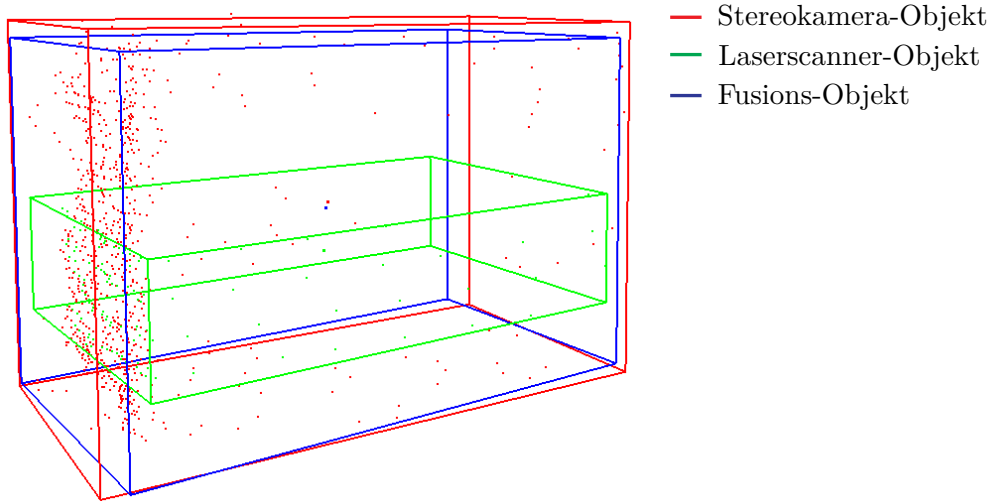


Abbildung 6.8: Darstellung der beiden Sensorobjekte (rot, grün) und dem daraus gebildeten Fusionsobjekt (blau) für das Fahrzeug aus der Szene in Abbildung 3.4(a)

6.3 Fusion der Tracks

Die für jeden Sensor aufgrund von aufeinander folgenden Beobachtungen abgeleiteten Bewegungszustände der Objekte werden in der High-Level-Fusion zusammengeführt. Im Gegensatz zur Low- und Mid-Level-Fusion ist die komplette Datenprozessierung der drei Ebenen der High-Level-Fusion vorgelagert. Die einzelnen Tracks der Sensoren werden einer zentralen Fusions-einheit zugeführt und analog zur MLF für die anschließende Fusion entsprechend den Bewegungsverläufen realer Objekte einander zugeordnet. Nach der Detektion und Bereinigung der Tracking-Fehler werden die korrespondierenden Tracks fusioniert und den entsprechenden globalen *System-Tracks* hinzugefügt (siehe Abbildung 6.9).

6.3.1 Assoziierung der Tracks

Analog zur Objektdatenfusion müssen die detektierten Tracks der einzelnen Sensoren zur Fusion einander zugeordnet werden. Prinzipiell wird dabei eine Erweiterung der Clusterung aus der Objektdatenfusion angewendet (siehe Abschnitt 6.2.1). Hierbei werden für die Track-Zuordnung nicht nur die Objektzustände der aktuellen Track-Positionen sondern zusätzlich die historischen Objektinformationen der zurückliegenden Positionen mit einbezogen, um die Zuverlässigkeit der Assoziation zu stützen.

Ein aus der Clusterung resultierender *Track-Cluster* besteht aus dem aktuellen und den zurückliegenden zusammengefassten $n - 1$ Clusterobjekten der assoziierten Tracks der einzelnen Sensoren (*Sensor-Tracks*) (siehe Abbildung 6.10). Ein Clusterobjekt bestimmt sich hierbei aus den Objektzuständen der Sensor-Tracks an den gleichen Track-Positionen (Zeitpunkten). Der Aufbau eines Clusterobjekts und die Distanzbestimmung zwischen einem Clusterobjekt und einem Sensor-Track-Objekt ist dabei ebenfalls wie in Abschnitt 6.2.1 definiert.

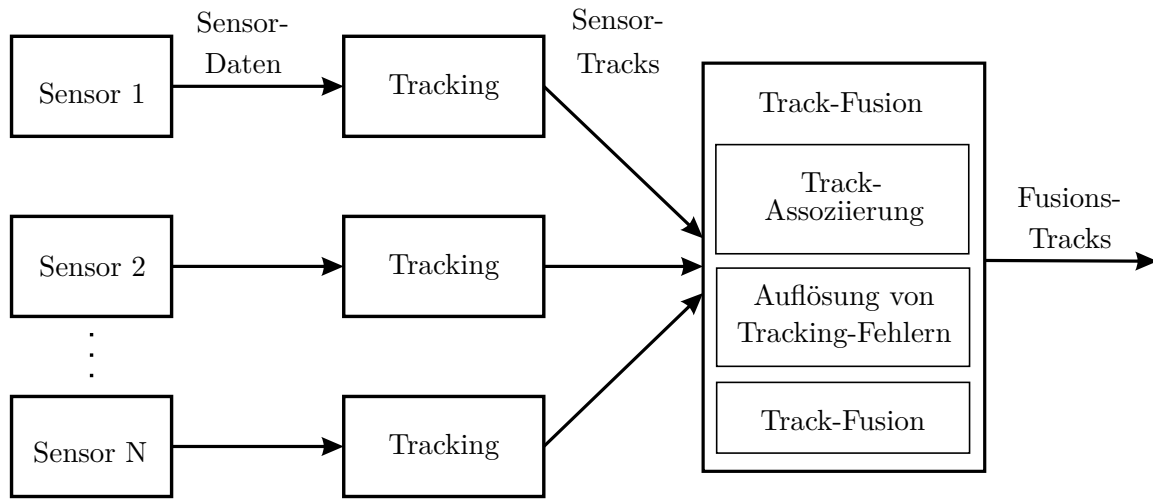


Abbildung 6.9: High-Level-Fusions-System

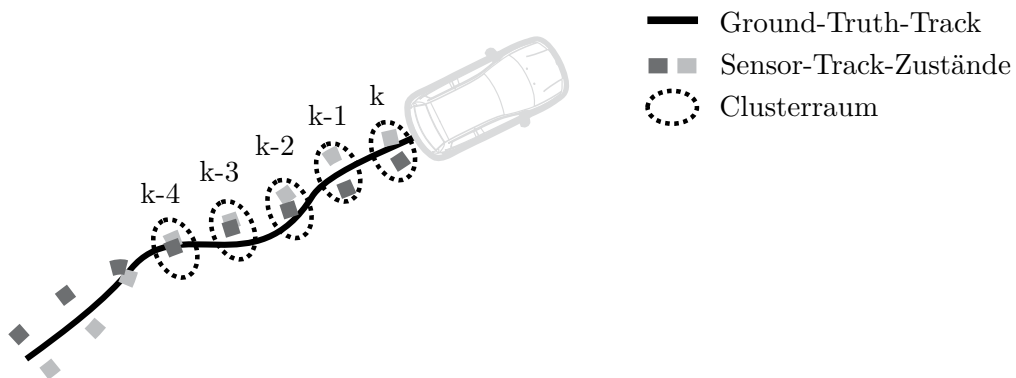


Abbildung 6.10: Zusammengehörige Objektzustände von Sensor-Tracks bilden einen Track-Cluster bei einer Assoziationstiefe von $n = 5$ Track-Positionen im Vergleich zum Ground-Truth-Track

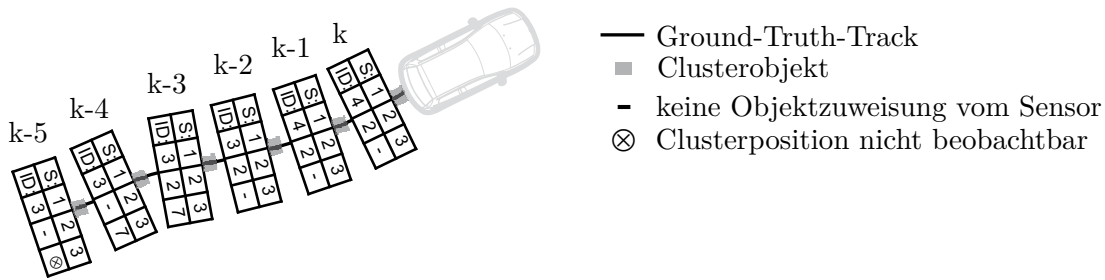


Abbildung 6.11: Darstellung eines System-Clusters, der aus dem aktuellen (k) und den vorherigen Track-Cluster-Objekten zusammengesetzt ist; jedes Track-Cluster-Objekt enthält Information über die Track-Zuordnungen aller Sensoren

Um eine Distanz d zwischen einem Track-Cluster und einem Sensor-Track zu bestimmen, werden die einzelnen Distanzen d_i zwischen den Cluster- und Track-Objekten der gleichen Position (Messzeit) mit dem gewichteten gleitenden Mittelwert (*Weighted Moving Average* – WMA) zusammengefasst:

$$d = \frac{nd_1 + (n-1)d_2 + \dots + 2d_{n-1} + d_n}{n + (n-1) + \dots + 2 + 1}. \quad (6.20)$$

Hierbei gehen die Distanzen zwischen den aktuelleren Objekten stärker gewichtet ein, was die Bestimmung der MOT-Effekte begünstigt (siehe Abschnitt 6.3.3). Die Assoziationsmatrix wird dann entsprechend Gleichung 6.1 für alle Objektpaarungen der aktuellen Track-Cluster-Liste und der Sensor-Track-Liste aufgestellt und analog zur Objektdaten-Clusterung aufgelöst. Dieser Vorgang wird wiederum für alle Sensor-Track-Listen durchgeführt und solange wiederholt, bis sich die Track-Cluster-Liste stabilisiert.

Zur Auflösung der Tracking-Effekte (siehe Abschnitt 5.2.3) werden die historischen Informationen zu den Zuordnungen der Sensor-Tracks zu den Track-Clustern benötigt. Diese werden in den *System-Clustern* gesammelt. Die Track-Cluster werden dazu in jedem Zeitschritt mit den System-Clustern nach dem oben beschriebenen Verfahren assoziiert, um anschließend die aktuellen Clusterobjekte der Track-Cluster den korrespondierenden System-Clustern hinzuzufügen (siehe Abbildung 6.11).

6.3.2 Auflösung von Fehl- und Falschdetektionen

Wie in Abbildung 6.1 dargestellt, werden fehlerhafte (FPT) oder fehlende (FNT) Tracks durch Fehler in den vorangegangenen Prozessstufen oder innerhalb der Messung verursacht. Zur Erkennung werden, wie in der Objektdatenfusion, die konkurrierenden Track-Informationen der beteiligten Sensoren herangezogen.

Analog zum Konflikt-Cluster bei der MLF zeigen die Konflikt-Track-Cluster (KTC) einen Widerspruch in der aktuellen Track-Detektion an. Ein KTC liegt vor, wenn nicht alle Sensoren, die den Track momentan beobachten können, diesen auch detektieren. Ein weiterer Widerspruch

stellt ein Wechsel einer Track-ID innerhalb eines System-Clusters in Folge einer wechselnden Objektverfolgung eines Tracks dar. Zur Auflösung werden die widersprüchlichen Sensoraussagen eines KTC gegeneinander abgewogen. Die in Abschnitt 6.2.2 definierten Mengen \mathbb{D} und $\bar{\mathbb{D}}$ beinhalten hier die Sensoren, die das zum Zeitpunkt k beobachtbare Objekt mit seinem zugehörigen Track-Cluster C detektieren bzw. nicht detektieren.

Die Konfliktsensoren der Mengen \mathbb{D} und $\bar{\mathbb{D}}$ werden durch die Gleichungen in 6.15 zur Bestimmung der Verbundwahrscheinlichkeiten $P(E_+|D_+^1, D_+^1, \dots, D_+^N)$ und $P(E_-|D_+^1, D_+^1, \dots, D_+^N)$ herangezogen. Hierbei entsprechen $P(E_+|D_+^1, D_+^1, \dots, D_+^N)$ und $P(E_-|D_+^1, D_+^1, \dots, D_+^N)$ den Wahrscheinlichkeiten, dass falls ein Track detektiert wurde, dieser existiert bzw. nicht existiert. Die Entscheidung über einen falsch detektierten Track-Cluster (FPT) oder fehlende Sensor-Tracks im Track-Cluster (FNT) kann somit über einen direkten Vergleich der beiden Wahrscheinlichkeiten mit

$$\begin{aligned} P(E_+|D_+^1, D_+^2, \dots, D_+^N) &> P(E_-|D_+^1, D_+^2, \dots, D_+^N) \rightarrow \text{FNT} \\ P(E_+|D_+^1, D_+^2, \dots, D_+^N) &< P(E_-|D_+^1, D_+^2, \dots, D_+^N) \rightarrow \text{FPT} \\ P(E_+|D_+^1, D_+^2, \dots, D_+^N) &= P(E_-|D_+^1, D_+^2, \dots, D_+^N) \rightarrow \text{Keine Entscheidung} \end{aligned} \quad (6.21)$$

getroffen werden.

Zur Bestimmung der benötigten Likelihoods $P(D_+|E_+)$ und $P(D_+|E_-)$ wird analog zur Objektdatenfusion die Segmentierungswahrscheinlichkeit SW der Sensoren für den Mittelpunkt des KTC M_C zum aktuellen Zeitpunkt k verwendet (siehe Abschnitt 4.1.3).

Da sich mit steigender Detektionszeit eines Tracks auch seine Existenzwahrscheinlichkeit erhöht, geht in die Bestimmung der SW nicht die analytisch bestimmbare Punktgenauigkeit für den Sensor am Punkt M_C sondern die aktuelle Zustandsgenauigkeit \mathbf{P}_k des Tracks ein. Längere Tracks besitzen daher über ihre abnehmende Zustandsvarianz kleinere F-Ellipsoide und somit eine höhere SW.

Die Standardabweichung $\bar{\sigma}$ zur Bestimmung der horizontalen SW (siehe Gleichung 4.4) wird demzufolge mit

$$\bar{\sigma} = \begin{cases} \sigma_{\mathbf{P}_x} & , \text{ mit } \bar{d} = \bar{d}_x \\ \sigma_{\mathbf{P}_y} & , \text{ mit } \bar{d} = \bar{d}_y \\ \sigma_{\mathbf{P}_z} & , \text{ mit } \bar{d} = \bar{d}_z \end{cases} \quad (6.22)$$

berechnet. Analog wird die vertikale bzw. die gesamte SW bestimmt (siehe Abschnitt 4.1.3).

Für die Sensoren der Menge $\bar{\mathbb{D}}$ ohne eine aktuelle Trackdetektion für C wird die SW an M_C ebenfalls in Abhängigkeit der Dauer der Nichtdetektion von C ermittelt. Die Anzahl n der zurückliegenden Trackingschritte ohne eine Track-Detektion bis zum aktuellen Zeitpunkt k lässt sich aus dem zu C zugehörigen System-Cluster ableiten. Wie in Abbildung 6.12 beispielhaft dargestellt, werden für den Sensor 3 ab dem Zeitpunkt $k - 2$ zum Track-Cluster keine Track-Detektionen mehr zugeordnet.

Um nun die Länge der Nichtdetektion ebenfalls über die Genauigkeit des Mittelpunktes des

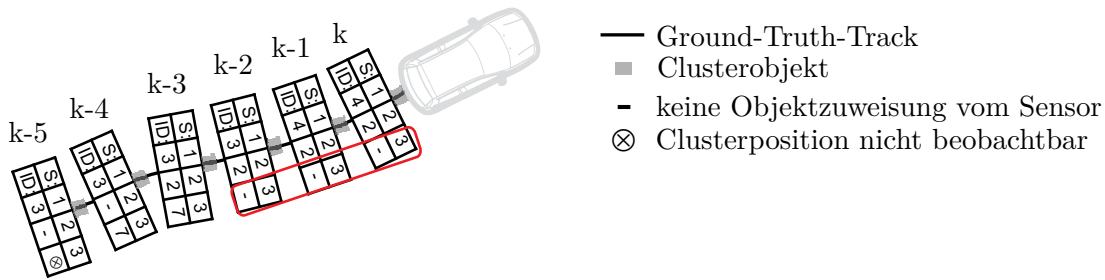


Abbildung 6.12: Erkennung eines FPT/FNT-Effekts für Sensor 3 und Bestimmung der zurückliegenden Trackingschritte ($n = 2$) ohne Track-Zuordnung zum Track-Cluster aus dem System-Cluster

Clusters M_C in die SW $P(S_{M_C}^X)$ einfließen zu lassen, werden die Punkte $M_{C^{k-n}} - M_{C^k}$ einem vereinfachten Tracking unterzogen. Das in Abschnitt 5.2 vorgestellte Objekt-Tracking wird durch die Reduzierung des Zustandsvektors \mathbf{x} um die Objektdimensionen zum Punkt-Tracking mit

$$\tilde{\mathbf{X}}_k = \begin{bmatrix} x_k & y_k & z_k & \psi_k & \omega_k & v_k \end{bmatrix}^T \quad (6.23)$$

umformuliert. In den Messwertvektor \mathbf{Z} gehen nunmehr sukzessive die 3D-Positionen von $M_{C^{k-n}} - M_{C^k}$ als Messwerte ein. Die Objektdimensionen und die Objektausrichtung entfallen:

$$\tilde{\mathbf{Z}}_k = \begin{bmatrix} x_k & y_k & z_k \end{bmatrix}^T \quad (6.24)$$

Die Gleichungen 5.14 - 5.23 vereinfachen sich dementsprechend.

Die aus dem Punkt-Tracking resultierende Positionsvarianz für M_C wird anschließend zur Bestimmung der SW $P(S_{M_C}^X)$ verwendet (siehe Gleichung 4.6).

Analog zur Auflösung der FPO- und FNO-Effekte in der Mid-Level-Fusion wird bei einem erkannten Falsch-Track (FPT) der entsprechende Track-Cluster aus der Clusterliste entfernt. Bei fehlenden Track-Zuordnungen bleibt der Track-Cluster unverändert bestehen.

6.3.3 Auflösung von Multi-Objekt-Tracks

Der durch eine falsche oder vertauschte Assoziierung hervorgerufene Multi-Objekt-Track (MOT) verfolgt innerhalb seines Lebenszyklusses im Gegensatz zum Einzel-Objekt-Track (EOT) fälschlicherweise mindestens zwei reale Objekte (siehe Abschnitt 5.2.3). Dieser Objektwechsel innerhalb eines Sensor-Tracks lässt sich anhand der zurückliegenden Track-Clusterungen des korrespondierenden System-Clusters wiederum nachvollziehen. Ändert sich die Track-ID eines Sensor-Tracks innerhalb eines System-Clusters, so wird von einem MOT-Effekt ausgegangen.

Abbildung 6.13 verdeutlicht dies anhand eines System-Clusters, der den Clusterungsverlauf dreier Sensor-Tracks verfolgt. Sensor 1 verfolgt hierbei den System-Track anfänglich mit Sensor-Track 3 und ordnet anschließend den Sensor-Track 4 zu.

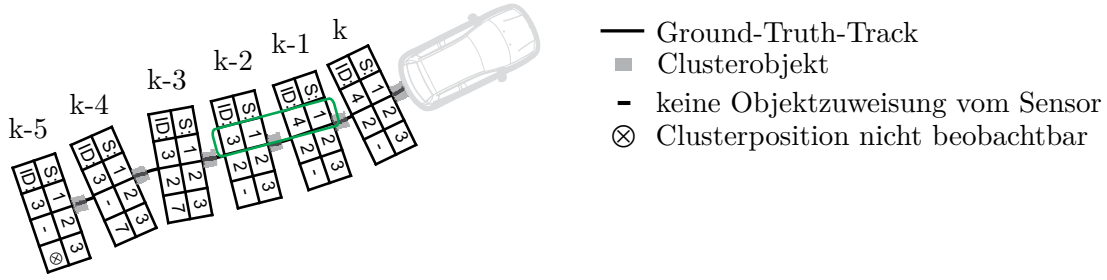


Abbildung 6.13: Erkennung eines MOT-Effekts für Sensor 1 durch wechselnde Track-ID zum Zeitpunkt $k - 1$ innerhalb des System-Clusters

Durch die Verwendung des WMA zur Assoziation zwischen einem Track-Cluster C_T und System-Cluster C_S (siehe Abschnitt 6.3.1) soll eine Falschzuordnung nach einem MOT-Effekt bei C_T vermieden werden. Durch die stärkere Gewichtung der aktuelleren Track-Zustände wird der fehlerhafte Einfluss des falschen Trackverlaufes des MO-Tracks in C_T in der Assoziation reduziert.

Wird ein MOT-Effekt innerhalb eines System-Clusters aufgedeckt, so wird der entsprechende Sensor-Track aus dem Track-Cluster entfernt. Alternativ könnte man Multi-Objekt-Tracks auf den korrekten Teil-Track nach dem Objektwechsel reduzieren, um nur die fehlerhafte Track-Information zu entfernen.

Im Zuge der Auswertung fiel bei der Bestimmung des System-Clusters zudem auf, dass ein Sensor-Track, verglichen mit zwei Sensor-Tracks mit halben Zustandsgenauigkeiten, den gleichen Einfluss bei der Bildung des Cluster-Tracks besitzt. Somit könnte ein MO-Track von zwei nur halb so genauen EO-Tracks über den damit bestimmten System-Cluster nicht erkannt werden. Um diesen möglicherweise unerwünschten Einfluss der Track-Genauigkeit im Vergleich zur Hypothesenanzahl zu reduzieren, wurde ein Gewichtungsfaktor f eingeführt. Die Gewichtung wird durch eine temporäre Anhebung aller clusterbildenden Kovarianzen p_{ij}^n der N beteiligten Sensoren erzielt. Hierbei werden die Abstände der einzelnen Kovarianzen zu ihrem gemeinsamen Mittelwert μ_P mit

$$\mu_P = \frac{1}{36N} \sum_{n=1}^N \sum_{i=1}^6 \sum_{j=1}^6 p_{ij}^n \quad (6.25)$$

um den Faktor f reduziert (gestaucht):

$$p_{ij}^n = (p_{ij}^n - \mu_P)f + \mu_P, \text{ mit } f \in [0, 1]. \quad (6.26)$$

Für $f = 1$ bleiben die Kovarianzen dementsprechend unverändert und für $f = 0$ werden alle Kovarianzen dem gemeinsamen Mittelwert μ_P gleichgesetzt. Durch eine Gleichsetzung aller Kovarianzen würde so eine mehrfache Track-Hypothese in der Clusterbildung immer eine höhere Gewichtung besitzen als eine einfache Track-Hypothese beliebig hoher Genauigkeit.

6.3.4 Zustandsfusion

Die zeitlich aufeinander folgenden Daten eines Sensor-Tracks sind über ihre gemeinsame historische Information zeitlich korreliert (Intra-Track-Korrelation). Zusätzlich besteht zwischen den korrespondierenden Sensor-Tracks verschiedener Sensoren eine gegenseitige Korrelation, die aufgrund gemeinsamer äußerer Einflüsse und Modellannahmen entstehen (Inter-Track-Korrelation) [10]. Bei einer Zustandsfusion der Sensor-Tracks durch einen Kalman-Filter können hierdurch, aufgrund der Tiefpass-Eigenschaft des Filters, Phasenverzögerungen entstehen [74].

Um die gemeinsame Information innerhalb eines Tracks und zwischen den Tracks zu filtern und so eine Verschiebung innerhalb der Gewichtung in der Track-Fusion zu vermeiden, wurden diverse Verfahren entwickelt. Im Folgenden werden drei bekannte Verfahren zur Track-Fusion vorgestellt und in einer anschließenden Messauswertung miteinander verglichen.

Bei einer Vernachlässigung der Inter-Track-Korrelationen zwischen den Tracks können die aktuellen Zustände \mathbf{x} und deren Kovarianzen \mathbf{P} über eine *Simple Convex Fusion* – SCF oder *Basic Convex Fusion* [78, 23], analog zur Mid-Level-Fusion, zusammengeführt werden:

Simple Convex Fusion

$$\mathbf{P}_k = \left[\sum_{i=1}^N (\mathbf{P}_{i_k})^{-1} \right]^{-1} \quad (6.27)$$

$$\mathbf{x}_k = \mathbf{P}_k \sum_{i=1}^N (\mathbf{P}_{i_k})^{-1} \mathbf{x}_{i_k} . \quad (6.28)$$

Ein Verfahren zur Bestimmung der Inter-Track-Korrelationen (Kreuz-Korrelationen) wurde in [4] vorgestellt, wodurch der SCF zum *Linear Combination with Cross Covariance* erweitert wurde [7]. Für die Bestimmung der Kreuz-Kovarianzen werden jedoch die historischen Parameter der vorangegangenen Kalman-Filter (\mathbf{K} , \mathbf{H} , \mathbf{A} , \mathbf{Q}) benötigt. In [23] wird daher angemerkt, dass es sinnvoller wäre, gleich die unkorrelierten Messdaten mit zu übertragen und zu fusionieren.

Covariance Intersection Fusion

Ein weiteres Verfahren, dass die korrespondierenden Track-Zustände und Kovarianzen über eine gewichtete konvexe Kombination fusioniert, ist die *Covariance Intersection Fusion* – CIF [108, 60]:

$$\mathbf{P}_k = \left(\omega (\mathbf{P}_{1_k})^{-1} + (1 - \omega) (\mathbf{P}_{2_k})^{-1} \right)^{-1} \quad (6.29)$$

$$\mathbf{x}_k = \mathbf{P}_k \left(\omega (\mathbf{P}_{1_k})^{-1} \mathbf{x}_{1_k} + (1 - \omega) (\mathbf{P}_{2_k})^{-1} \mathbf{x}_{2_k} \right), \text{ mit } \omega \in [0, 1] . \quad (6.30)$$

Zur Bestimmung der Gewichtung ist hierbei vorteilhafterweise die Kenntnis über gegenseitige Korrelationen nicht notwendig. Das Gewicht ω wird dabei zumeist über eine Minimierung der Determinante von \mathbf{P} bestimmt, wobei für alle $\omega \in [0, 1]$ die Konsistenz der Zustandsschätzung

garantiert ist. Um die CIF für mehr als zwei Sensoren anzuwenden, kann folgende allgemeine Form angegeben werden:

$$\mathbf{P}_k = \left[\sum_{i=1}^N \omega_i (\mathbf{P}_{i_k})^{-1} \right]^{-1} \quad (6.31)$$

$$\mathbf{X}_k = \mathbf{P}_k \sum_{i=1}^N \omega_i (\mathbf{P}_{i_k})^{-1} \mathbf{x}_{i_k} \quad (6.32)$$

$$\sum_{i=1}^N \omega_i = 1 . \quad (6.33)$$

Fast Covariance Intersection Fusion

Um die hierfür komplexe und teilweise numerisch instabile Optimierung zur Bestimmung der Gewichte ω_i für mehr als zwei Sensoren zu umgehen, wird in [38] folgende Vereinfachung als *Fast Covariance Intersection Fusion* – FCIF vorgeschlagen:

$$\omega_i = \frac{\det[(\tilde{\mathbf{P}}_k)^{-1}] + \det[(\mathbf{P}_{i_k})^{-1}] - \det[(\tilde{\mathbf{P}}_k)^{-1} - (\mathbf{P}_{i_k})^{-1}]}{N \det[(\tilde{\mathbf{P}}_k)^{-1}] + \sum_{j=1}^N [\det[(\mathbf{P}_{j_k})^{-1}] - \det[(\tilde{\mathbf{P}}_k)^{-1} - (\mathbf{P}_{j_k})^{-1}]]} \quad (6.34)$$

$$(\tilde{\mathbf{P}}_k)^{-1} = \sum_{i=1}^N (\mathbf{P}_{i_k})^{-1} . \quad (6.35)$$

Für den Fall, dass die zu fusionierenden Track-Zustände eine statistische Inkonsistenz aufweisen, z.B. die Distanzen zwischen den Zuständen nicht innerhalb der Wahrscheinlichkeitsbereiche ihrer Kovarianzen liegen, wurde die CI um die *Covariance Union* ergänzt [109]. Die Schwierigkeit besteht jedoch in der Bestimmung eines Schwellwertes für den optimalen Wechsel zwischen CI und CU [111] sowie einer echtzeitfähigen Implementierung für mehrere Sensoren.

Information Fusion

Die *Information Fusion* – IF oder auch *Information De-correlation* leitet sich aus der kovarianzinversen Darstellung des Kalman-Filters als Informationsfilter ab. Vorgestellt in [22] und [24], versucht die IF die gemeinsamen Informationen der aufeinander folgenden Zustände eines Sensor-Tracks und zwischen den Sensor-Tracks zu bestimmen und von der Fusion zu subtrahieren:

$$\mathbf{P}_k = \left[(\mathbf{P}_k^-)^{-1} + \sum_{i=1}^N \left[(\mathbf{P}_{i_k})^{-1} - (\mathbf{P}_{i_k}^-)^{-1} \right] \right]^{-1} \quad (6.36)$$

$$\mathbf{x}_k = \mathbf{P}_k \left[(\mathbf{P}_k^-)^{-1} \mathbf{x}_k^- + \sum_{i=1}^N \left[(\mathbf{P}_{i_k})^{-1} \mathbf{x}_{i_k} - (\mathbf{P}_{i_k}^-)^{-1} \mathbf{x}_{i_k}^- \right] \right]. \quad (6.37)$$

Der Ansatz wird als optimal betrachtet, falls kein Prozessrauschen vorhanden ist oder die angebundenen Sensoren regelmäßig in einem gemeinsamen Intervall ihre Daten senden (wie es in dieser Arbeit der Fall ist) [19]. Ein weiterer Vorteil ist die Anbindung beliebig vieler Sensoren sowie die vergleichsweise einfache Implementierung und Berechnung. Es werden keine zusätzlichen Informationen benötigt, da die Zustände und Kovarianzen der Sensoren und des System-Tracks aus dem vorherigen Schritt nur zwischengespeichert werden müssen [23].

Nach der Fusion der Zustände aller zusammengehörigen Sensor-Tracks eines Track-Clusters wird, außer bei der IF, der zugehörige System-Track um den resultierenden Zustand und dessen Kovarianzmatrix erweitert. Beim IF wird der zugehörige System-Track bereits innerhalb der Fusion erweitert. Falls keine Systemzuordnung existiert, wird ein neuer System-Track initialisiert. Für den Fall, dass einem System-Track für eine bestimmte Anzahl in Folge kein Track-Cluster zugeordnet wurde, wird dieser entfernt.

6.4 Kombination der Fusionsebenen

Betrachtet man rückblickend die Vor- und Nachteile der drei Fusionsebenen, so lässt sich feststellen, dass mit fortschreitender Verarbeitung und Fusion der Sensordaten die Genauigkeit der Objektbestimmung abnimmt (Auswertungsergebnisse siehe Abschnitt 7.6).

Andererseits lassen sich zu einem in der Datenprozessierung späteren Fusionszeitpunkt die Fehler der vorangegangenen Prozessschritte besser detektieren und entfernen.

Die logische Konsequenz liegt in der Bündelung der spezifischen Vorteile der drei Fusionsebenen durch eine sequentielle Kombination der Verfahren. Hierbei werden im Sinne eines Abstiegs die Objektdaten, die durch Fusion auf den oberen Verarbeitungsebenen gewonnen wurden, auf die untere Punktdatenebene projiziert, um anschließend in einem erneuten Aufstieg in der Prozesskette eine optimale Objektbeschreibung zu erhalten (siehe rot dargestellter Datenfluss in Abbildung 6.14).

Im ersten Schritt werden hierfür die Sensordaten durch die MLF und HLF separat nach den in den Abschnitten 6.2 und 6.3 vorgestellten Verfahren parallel fusioniert. Anschließend werden die resultierenden Objektdaten der HLF mit den Objektdaten aus der MLF assoziiert (siehe Abschnitt 5.1). Besteht ein Konflikt im Track-Cluster eines HLF-Objektes so kann das korrespondierende MLF-Objekt zur Verifizierung der HLF-Entscheidung herangezogen werden, um die HLF-Entscheidung ggf. zu revidieren. Hierbei wird auf der Objektebene nach den Ursachen für einen detektierten Falsch- oder Fehl-Track gesucht, die wie in Abbildung 6.1 dargestellt, auf ein Falsch- oder Fehl-Objekt, einem Zusammenfassungs- oder Fragmentierungs-Objekt oder einem Assoziierungsfehler in der Track-Clusterung zurückzuführen sind. Bei einem zugeordne-

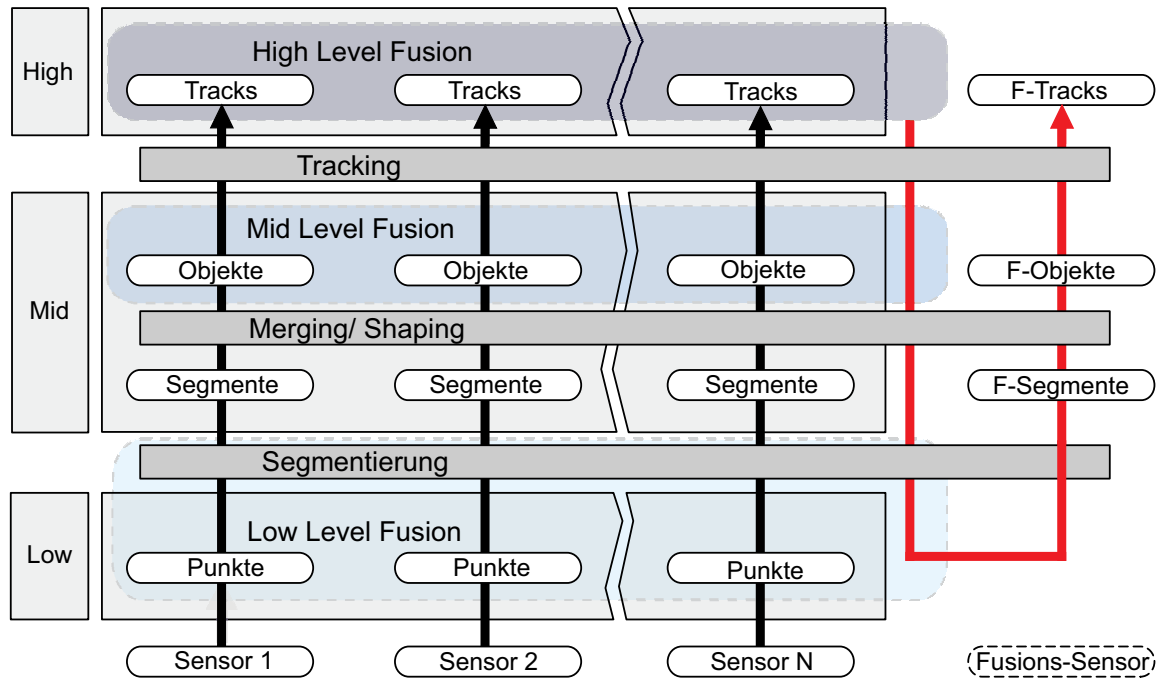


Abbildung 6.14: Sequentielle Kombination der drei Fusionsverfahren auf den unterschiedlichen Ebenen

tem HLF- und MLF-Cluster-Paar unterstützen hierbei die Hypothesen für ein Falsch- bzw. Fragmentierungs-Objekt die Hypothese eines Falsch-Tracks. Analog unterstützen die Hypothesen für ein Fehl- bzw. Zusammenfassungs-Objekt die Hypothese eines Fehl-Tracks.

Die Menge der verifizierten HLF-Objekte $\check{\check{O}}$ werden hiernach in den zusammengefassten Raum der räumlichen Punktdaten der Sensoren übertragen. Für jedes Objekt \check{O} aus $\check{\check{O}}$ wird eine separate und lokal segmentierende Fusion durchgeführt. Der Segmentierungsraum wird hierbei durch die zweifach vergrößerte Objektbox $\check{\check{O}}$ von \check{O} bestimmt, um die Ungenauigkeit der Objektbeschreibung von \check{O} durch die Fusion auf den höheren Ebenen zu kompensieren:

$$\begin{aligned} d_{x_{\check{\check{O}}}} &= 2d_{x_{\check{O}}} \\ d_{y_{\check{\check{O}}}} &= 2d_{y_{\check{O}}} \\ d_{z_{\check{\check{O}}}} &= 2d_{z_{\check{O}}} . \end{aligned} \quad (6.38)$$

In Tabelle 6.3 sind die entsprechenden Maßnahmen zur Vermeidung der Tracking- und Segmentierungsfehler (siehe Abschnitt 4.2.4) in der lokalen Segmentierung aufgeführt.

Die resultierenden Segmente eines Ausgangsobjekts \check{O} werden anschließend zusammengefasst und über das Shaping-Verfahren die Objektdrehung geschätzt (siehe Abschnitt 4.2.2). Das dementsprechend bestimmte Objekt \hat{O} besitzt somit die gewünschte Genauigkeit der LLF bei

Fehlertyp	Maßnahmen zur Fehlervermeidung
Falschdetektion (FPO, FPT)	Keine weiteren Maßnahmen notwendig, da falsche Objekt/Track-Hypothesen aus der Hypothesenmenge durch MLF und HLF bereits entfernt
Fehldetektion (FNO, FNT)	Keine weiteren Maßnahmen notwendig, da Fehldetektionen durch Objektdetektionen anderer Sensoren kompensiert werden
Zusammenfassungs-Objekt (ZO)	Im Falle einer detektierten Objektzusammenfassung (ZO) für den zugehörigen Cluster von \check{O} gehen zur Vermeidung nur die Punktdaten der Sensoren in die lokale Segmentierung ein, die bei der Konfliktauflösung als korrekt detektierend bestimmt wurden
Fragmentierungs-Objekt (FO)	Im Falle einer detektierten Objektfragmentierung (FO) für den zugehörigen Cluster von \check{O} sind keine weiteren Maßnahmen notwendig, da die Punktinformationen weiterer Sensoren eine Fragmentierung in der lokalen Segmentierung an der Stelle von \check{O} vermeiden

Tabelle 6.3: Maßnahmen zur Fehlervermeidung in der lokalen Segmentierung

gleichzeitiger Vermeidung der Segmentierungsfehler.

Da der Bezug zum bereits getrackten Objekt \check{O} und somit zum zugeordneten System-Track über die Track-ID besteht, kann auf eine Assoziierung für das anschließende Tracking bei \hat{O} verzichtet werden. Fehler, die auf die Assoziierung zurückzuführen sind (FNZ, FPZ), können somit ebenfalls vermieden werden (siehe Abschnitt 5.2.3).

Alternativ sind noch weitere Kombinationsmöglichkeiten der Fusionsebenen denkbar. Die Verifikation der HLF durch die MLF könnte entfallen, um die LLF direkt durch die HLF-Objekte zu steuern.

Liefern die Sensoren keine Punktdaten, könnten die HLF-Objektdaten auf die MLF-Objektdaten abgebildet werden, um eine genauere Objektbeschreibung zu erzielen und die Fehlererkennung zu stützen.

6.5 Zusammenfassung

Im Kapitel zur Multilevel-Datenfusion wurden die unterschiedlichen Verfahren zur Fusion der Sensordaten auf den drei Verarbeitungsebenen vorgestellt. Die Punktdatenfusion auf der untersten Ebene fasst die Punktdaten in einer gemeinsamen Segmentierung zu Objektinformationen zusammen. Objektdaten der mittleren Ebene werden in der Objektdatenfusion über einen mehrstufigen Prozess fusioniert. Durch die Objektclusterung zur Zuordnung der Sensorobjekte lassen sich zusätzlich Segmentierungseffekte erkennen. So können sowohl Objektfragmentierungen und -zusammenfassungen als auch Fehl- und Falschdetektionen erfasst und bereinigt werden. Die

Grundlage zur Konfliktauflösung stellen dabei die Segmentierungs- und Fehlermodelle der Sensoren dar. Analog werden die Track-Daten der oberen Ebene in der High-Level-Fusion vereint. Nach der Clusterung der Tracks unter Einbeziehung der historischen Daten, dienen die Track-Cluster zur Aufdeckung und Bereinigung der Tracking-Effekte wie Fehl- und Falsch-Tracks sowie Multi-Objekt-Tracks. Schließlich wurde eine Erweiterung des MMFF zur Kopplung der Fusionsebenen eingeführt, um die Vorteile der Fehlerbereinigung der oberen mit der optimalen Objektbestimmung der unteren Ebenen zu vereinen.

7 Experimentelle Ergebnisse

Die in den vorangegangenen Kapiteln vorgestellten und neu entwickelten Methoden sollen im Folgenden durch gemessene oder synthetisch erzeugte Daten überprüft und bewertet werden. Die für alle Experimente gleichermaßen verwendeten Parameter der Sensormodelle, der Kalibriermodelle sowie des Tracking-Filters sind im Anhang entsprechend dokumentiert (siehe Anhang-Abschnitt 2).

7.1 Messgenauigkeiten der Sensoren

Eine wesentliche Grundlage der folgenden Messungen und Simulationen bilden die Genauigkeitsangaben der Punktmessungen der beiden verwendeten Sensoren. Diese leiten sich aus dem im Vorfeld bestimmten Messrauschen und den Fehlermodellen der Sensoren ab (siehe Abschnitt 3.3).

Zum anschaulichen Verständnis und zum Nachvollziehen der weiteren Messergebnisse wurden die Messgenauigkeiten der Sensoren jeweils für eine Koordinatenrichtung der drei SKS-Richtungen in den Abbildungen 7.1 - 7.3 im jeweiligen messbaren Sensorbereich beispielhaft dargestellt. Die analytisch bestimmten Genauigkeiten entsprechen hierbei den Punktmessungen in einem gitterartigen Messfeld von 51 m Breite und Tiefe. Die betreffenden Messpunkte sind in 1 m Höhe mit einem Abstand von 1 m zu den Nachbarpunkten in X- und Y-Richtung im Gitter angeordnet. Das Messfeld erstreckt sich entlang der X-Achse des SKS im Bereich zwischen 5 m und 55 m und entlang der Y-Achse im Bereich zwischen -25 m und 25 m, um den für die Verkehrserfassung interessanten Messbereich exemplarisch abzudecken.

7 Experimentelle Ergebnisse

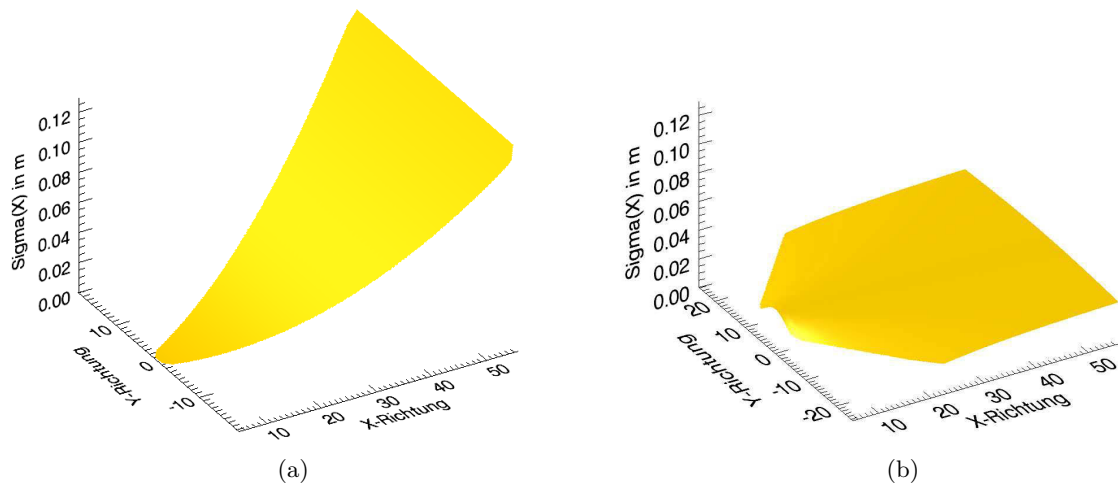


Abbildung 7.1: Berechnete Standardabweichung der Messpunkte in X-Richtung des SKS für das Stereokamerasystem (a) und den Laserscanner (b)

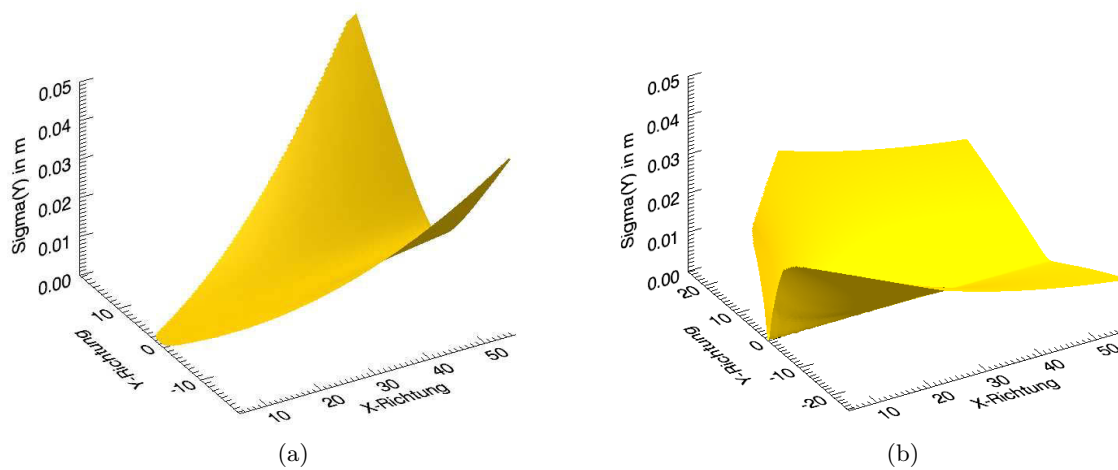


Abbildung 7.2: Berechnete Standardabweichung der Messpunkte in Y-Richtung des SKS für das Stereokamerasystem (a) und den Laserscanner (b)

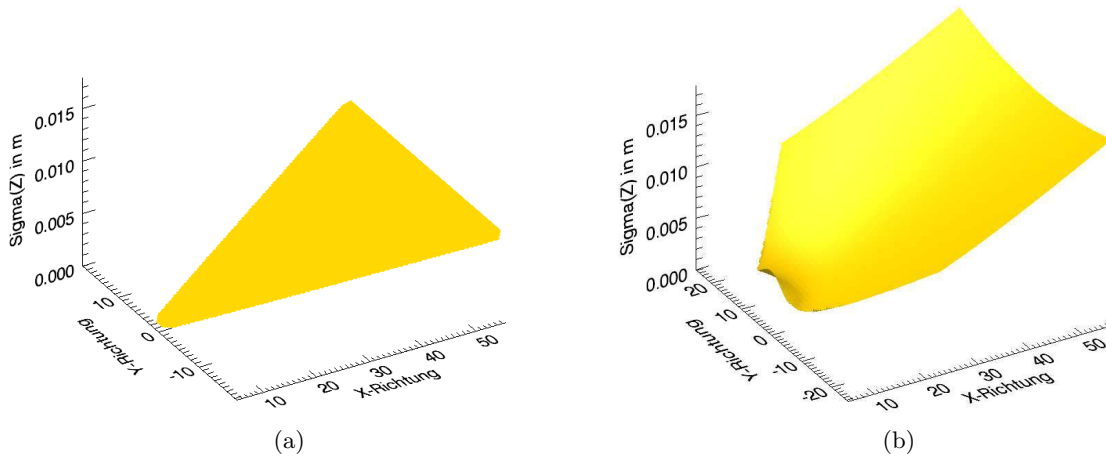


Abbildung 7.3: Berechnete Standardabweichung der Messpunkte in Z-Richtung des SKS für das Stereokamerasystem (a) und den Laserscanner (b)

7.2 Tracking mit Referenzpunktwechsel

Um eine Genauigkeitssteigerung in der Objektverfolgung durch einen dynamischen Wechsel des Referenzpunktes (RPT) im Vergleich zum Mittelpunkt-Tracking (MPT) zu demonstrieren (siehe Abschnitt 5.3.1), wurde eine Kreisfahrt für ein Testobjekt mit den Abmaßen $d_x = 4$ m, $d_y = 2$ m und $d_z = 2$ m simuliert. Das Testobjekt bewegt sich hierbei gleichmäßig mit 20 km/h auf einer Kreisbahn mit einem Radius $r = 10$ m um den Kreismittelpunkt K . Der Kreismittelpunkt liegt im SKS an der Position $K_x = 30$ m, $K_y = 0$ m und $K_z = 0$ m. Pro Simulation werden zehn Kreisfahrten in Folge durchgeführt.

Im **Szenario A** wird dabei angenommen, dass die wahren Objektausmaße und somit die wahre Position des Mittelpunktes idealerweise bekannt sind. In Abbildung 7.4 sind die Ergebnisse für ein Objekttracking mit unverrauschten, perfekten Messwerten durch das Verhältnis zwischen dem gemessenen Überlappungsvolumen V_U mit der Referenz zum kompletten Referenzvolumen V_R dargestellt. Wie ebenfalls der Tabelle 7.1 zu entnehmen ist, geht hierbei die durchschnittliche Abweichung von V_U zu V_R (RMS) beim MPT gegen Null. Für das RPT liegt die Objektabweichung zur Referenz für unverrauschte Messwerte im Schnitt bei 0.176 m^3 , was sich durch den Einfluss des EKF bei der Bestimmung des Referenzpunktdrehwinkels ψ_R und der Drehgeschwindigkeit ω_R erklären lässt.

Bei der Addition eines zufälligen Fehlers zu den Messwerten (sowohl Position als auch Drehwinkel) verstärken sich ebenfalls die durchschnittlichen Objektabweichungen für das MPT und das RPT, wobei der relative Unterschied zwischen MPT und RPT unabhängig von der Streustärke annähernd gleich bleibt (siehe Abbildungen 7.5 und 7.6). Für den Fall einer idealen aber unrealistischen dauerhaften Beobachtbarkeit von mindestens drei Referenzpunkten und der damit möglichen idealen Schätzung des Objektmittelpunktes wäre somit das MPT dem RPT leicht überlegen.

7 Experimentelle Ergebnisse

	$\sigma = 0$	$\sigma = 0.1$	$\sigma = 0.2$
RMS Mittelpunkt-Tracking (m^3)	0.031	0.309	0.563
RMS Referenzpunkt-Tracking (m^3)	0.176	2.182	1.658

Tabelle 7.1: Vergleich der mittleren Abweichung (RMS) der gemessenen Objekte zur Referenz zwischen dem Mittelpunkt- und Referenzpunkt-Tracking bei einer idealen Wahrnehmung der Objektausmaße; die simulierten Messungenauigkeiten beziehen sich auf die Lage (in m) und die Ausrichtung (in rad) des Objekts (Szenario A)

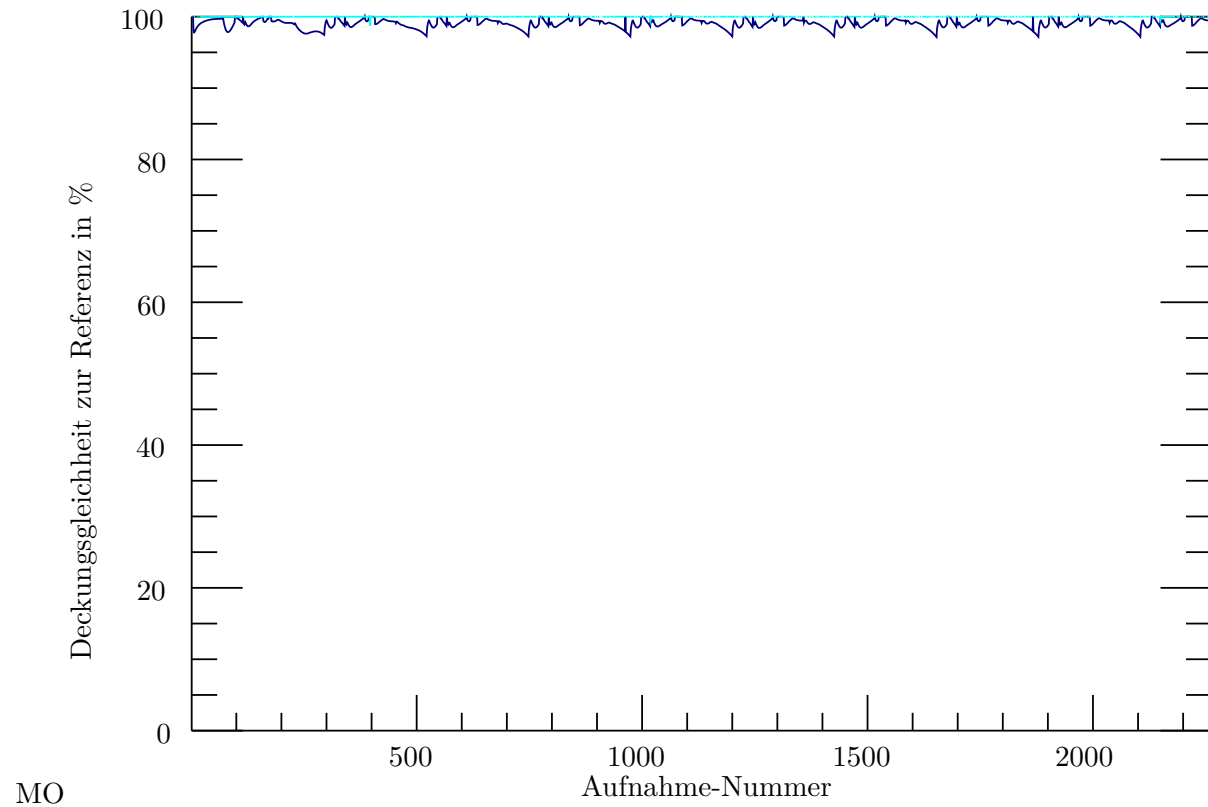


Abbildung 7.4: Verhältnis von Überlappungsvolumen V_U zum Referenzvolumen V_R für das Mittelpunkt-Tracking (hellblau) und Referenzpunkt-Tracking (dunkelblau) in einer simulierten idealen Objektsicht und einem nicht vorhandenen Messwertrauschen (Szenario A)

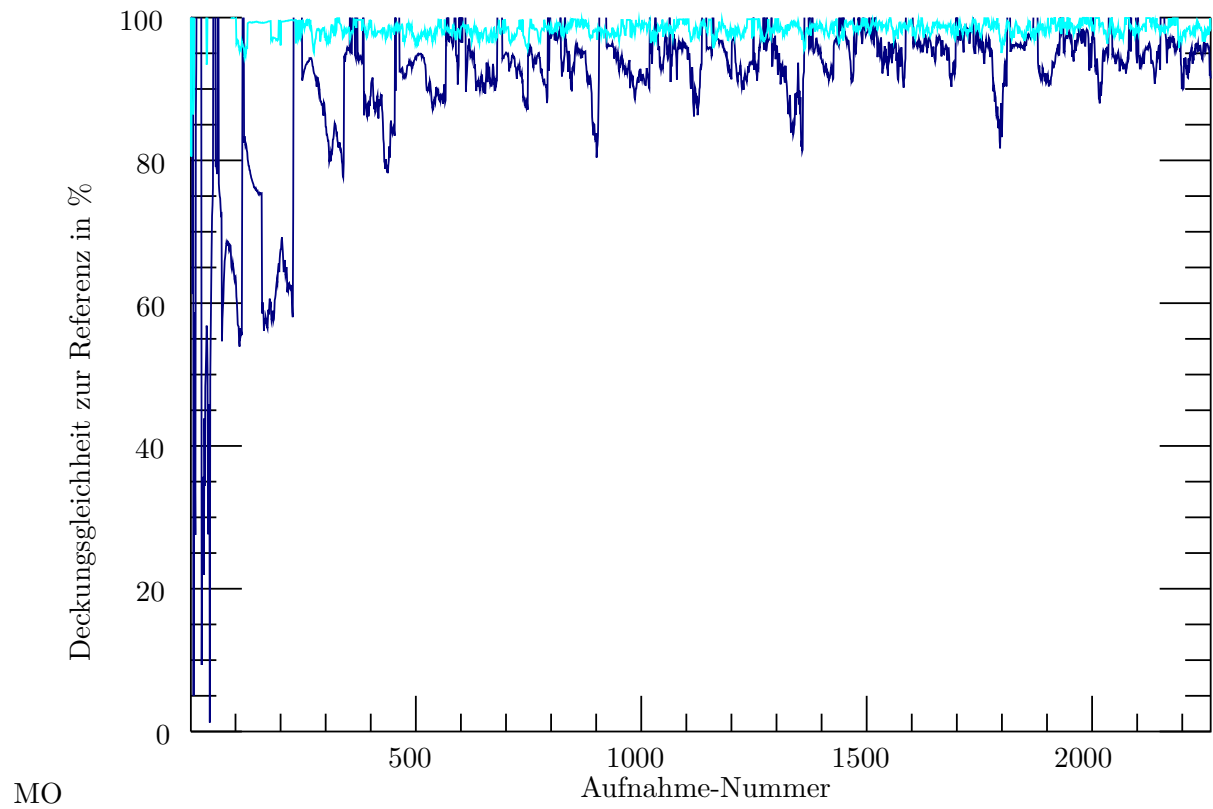


Abbildung 7.5: Verhältnis von Überlappungsvolumen V_U zum Referenzvolumen V_R für das Mittelpunkt-Tracking (hellblau) und Referenzpunkt-Tracking (dunkelblau) in einer simulierten idealen Objektsicht und einem Messwertrauschen von $\sigma = 0.1$ m bzw. rad (Szenario A)

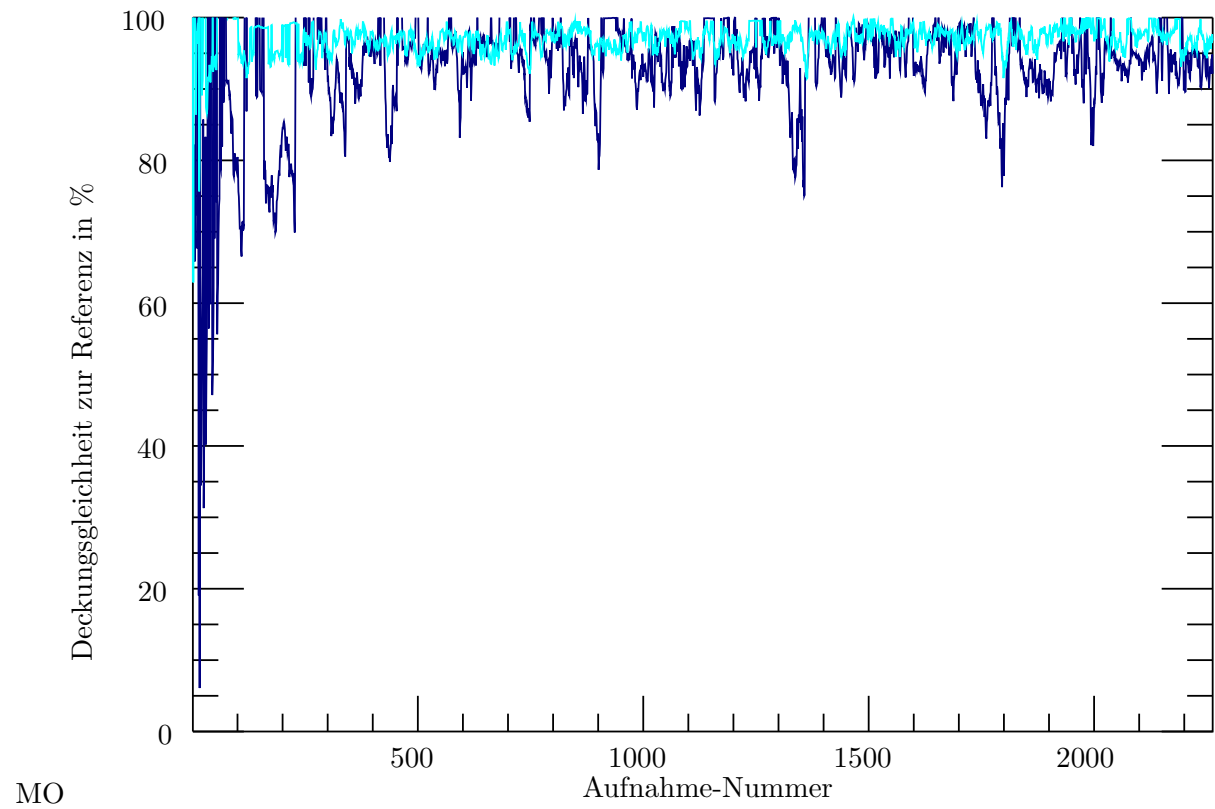


Abbildung 7.6: Verhältnis von Überlappungsvolumen V_U zum Referenzvolumen V_R für das Mittelpunkt-Tracking (hellblau) und Referenzpunkt-Tracking (dunkelblau) in einer simulierten idealen Objektsicht und einem Messwerttrauschen von $\sigma = 0.2$ m bzw. rad (Szenario A)

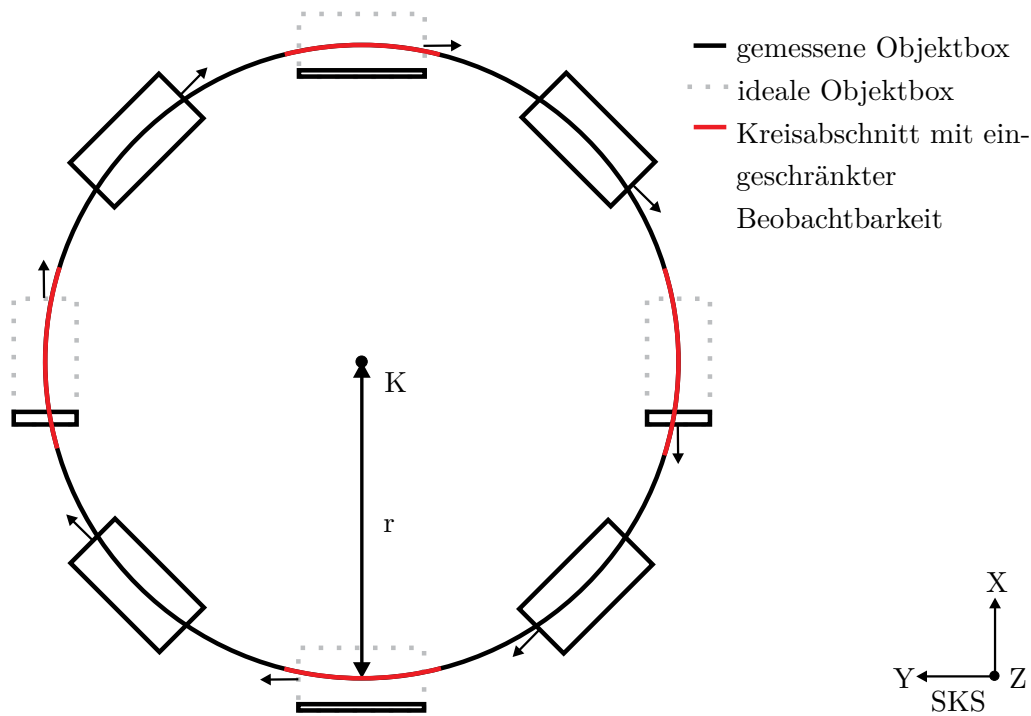


Abbildung 7.7: Simulation einer Objektkreisfahrt mit einer Reduktion der messbaren Objektausmaße d_x und d_y in Bereichen eingeschränkter Beobachtbarkeit vom Standpunkt der Sensoren (Szenario B)

Durch die wechselnde Lage eines zu verfolgenden Objekts im Bezug zum Sensorstandpunkt kann durch die partielle Eigenverdeckung des Objekts die Bestimmung des Objektmittelpunkts beeinträchtigt werden. Um dies in der Simulation zusätzlich abzubilden, werden im **Szenario B** die beobachtbaren und somit messbaren Objektausmaße entsprechend dem Standpunkt und dem Sichtbereich der Sensoren angepasst. Bei einer nahezu geradlinigen Objektbewegung in X-Richtung des SKS wird die gemessene Objektausdehnung d_x in X-Richtung des OKS um den Faktor 0.1 reduziert. Entsprechend wird im Bereich einer nahezu geradlinigen Bewegung in Y-Richtung des SKS die gemessene Objektausdehnung d_y in Y-Richtung des OKS ebenfalls um den Faktor 0.1 reduziert (siehe Abbildung 7.7). Die restlichen Parameter der Simulation bleiben unverändert.

Bei der idealisierten Simulation ohne Messrauschen (siehe Abbildung 7.8), als auch bei der Simulation mit Hinzunahme von Messrauschen (siehe Abbildungen 7.9 und 7.10) ist die durchschnittliche Objektabweichung zum Referenzobjekt beim RPT geringer als beim MPT (siehe Tabelle 7.2). Wie zu erwarten, kompensiert das RPT die wechselnden Messungen der Objektausmaße besser als das MPT. Das maximale Überlappungsvolumen beträgt nur ca. 70 %, da sich der EKF aufgrund der stetig wechselnden Objektdimensionen auf diesen Wert einpegelt. Die geringere Objektabweichung für das MPT bei einer Messwertstreuung von $\sigma = 0.2$ m im Vergleich zur nicht vorhandenen Messwertstreuung lässt sich durch das geringere Vertrauen in die

7 Experimentelle Ergebnisse

Messwerte und dem steigenden Vertrauen in das Systemmodell bei steigender Messwertstreuung im EKF erklären.

	$\sigma = 0$	$\sigma = 0.1$	$\sigma = 0.2$
RMS Mittelpunkt-Tracking (m^3)	6.135	6.337	5.666
RMS Referenzpunkt-Tracking (m^3)	4.505	5.046	4.955

Tabelle 7.2: Vergleich der mittleren Abweichung (RMS) der gemessenen Objekte zur Referenz zwischen dem Mittelpunkt- und Referenzpunkt-Tracking bei einer partiellen Einschränkung in der Beobachtbarkeit der Objektausmaße; die simulierte Streuung σ bezieht sich auf die Lage (in m) und die Ausrichtung (in rad) des Objekts (Szenario B)

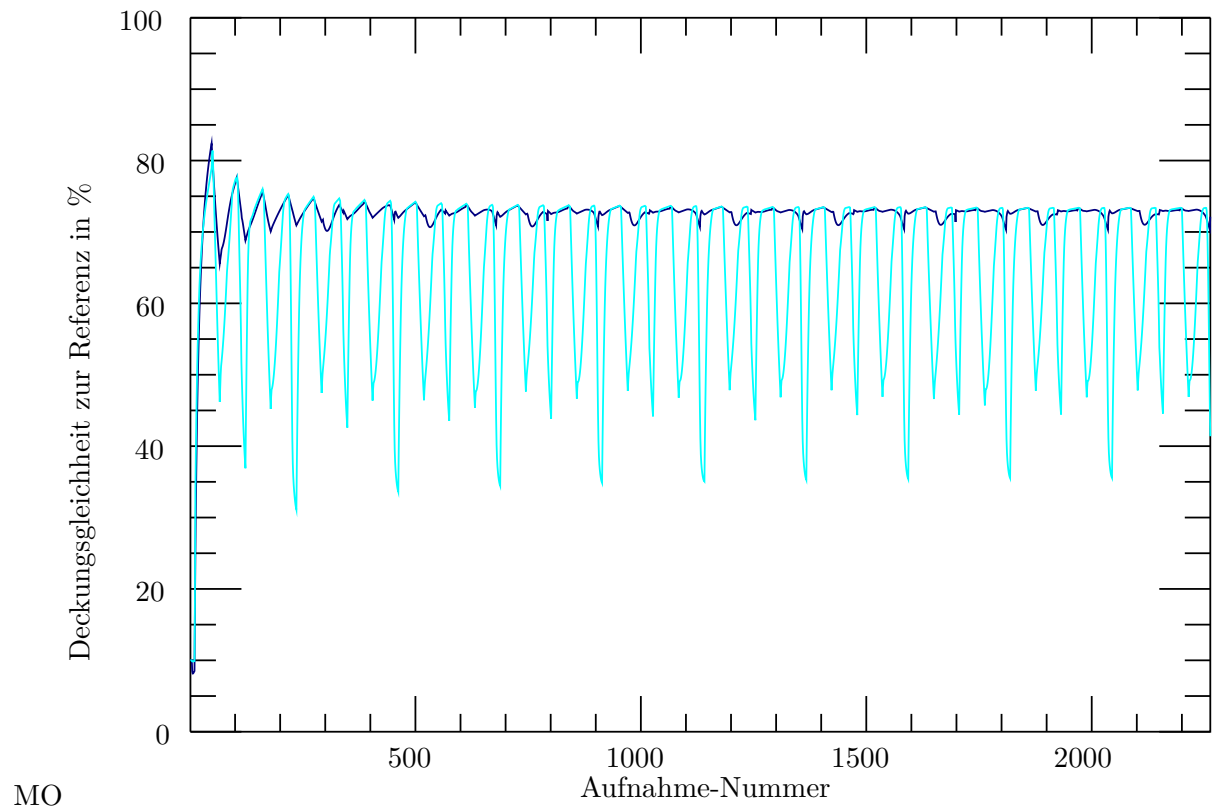


Abbildung 7.8: Verhältnis von Überlappungsvolumen V_U zum Referenzvolumen V_R für das Mittelpunkt-Tracking (hellblau) und Referenzpunkt-Tracking (dunkelblau) in einer simulierten partiell reduzierten Objektsicht und einem nicht vorhandenen Messwertrauschen (Szenario B)

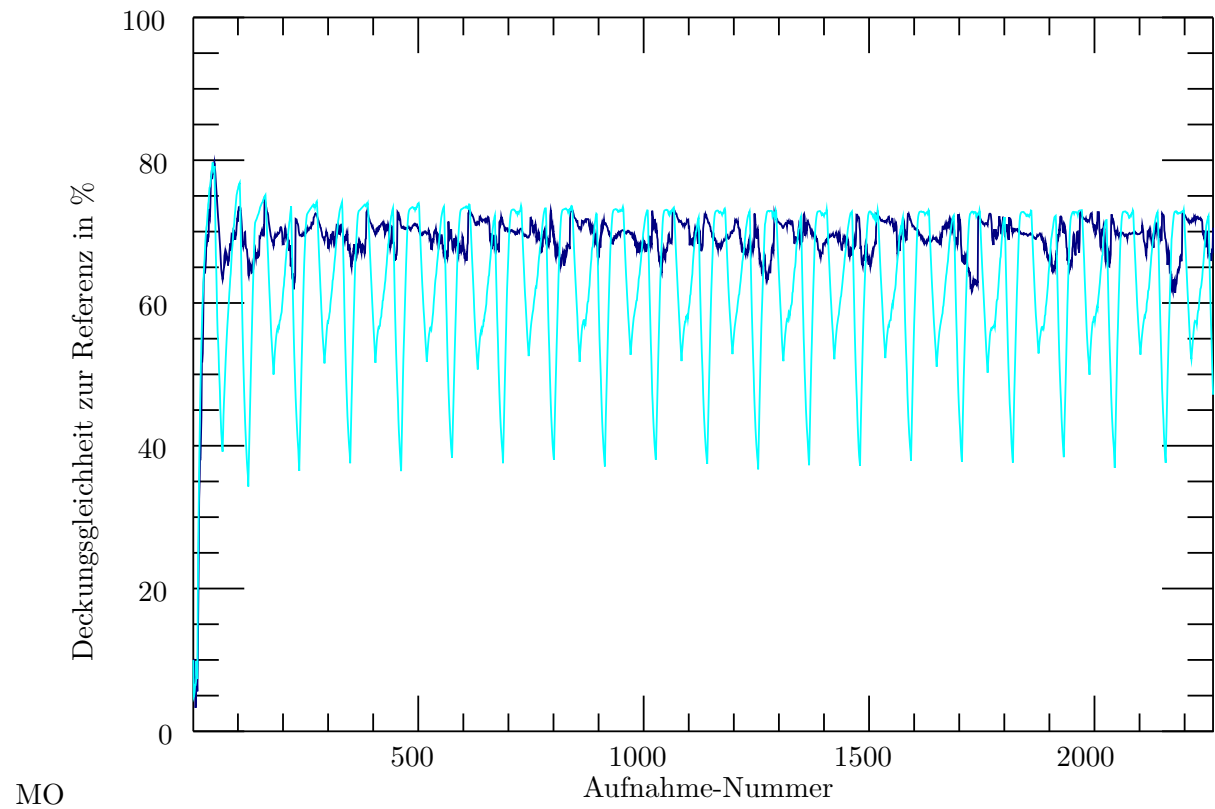


Abbildung 7.9: Verhältnis von Überlappungsvolumen V_U zum Referenzvolumen V_R für das Mittelpunkt-Tracking (hellblau) und Referenzpunkt-Tracking (dunkelblau) in einer simulierten partiell reduzierten Objektsicht und einem Messwertrauschen von $\sigma = 0.1$ m bzw. rad (Szenario B)

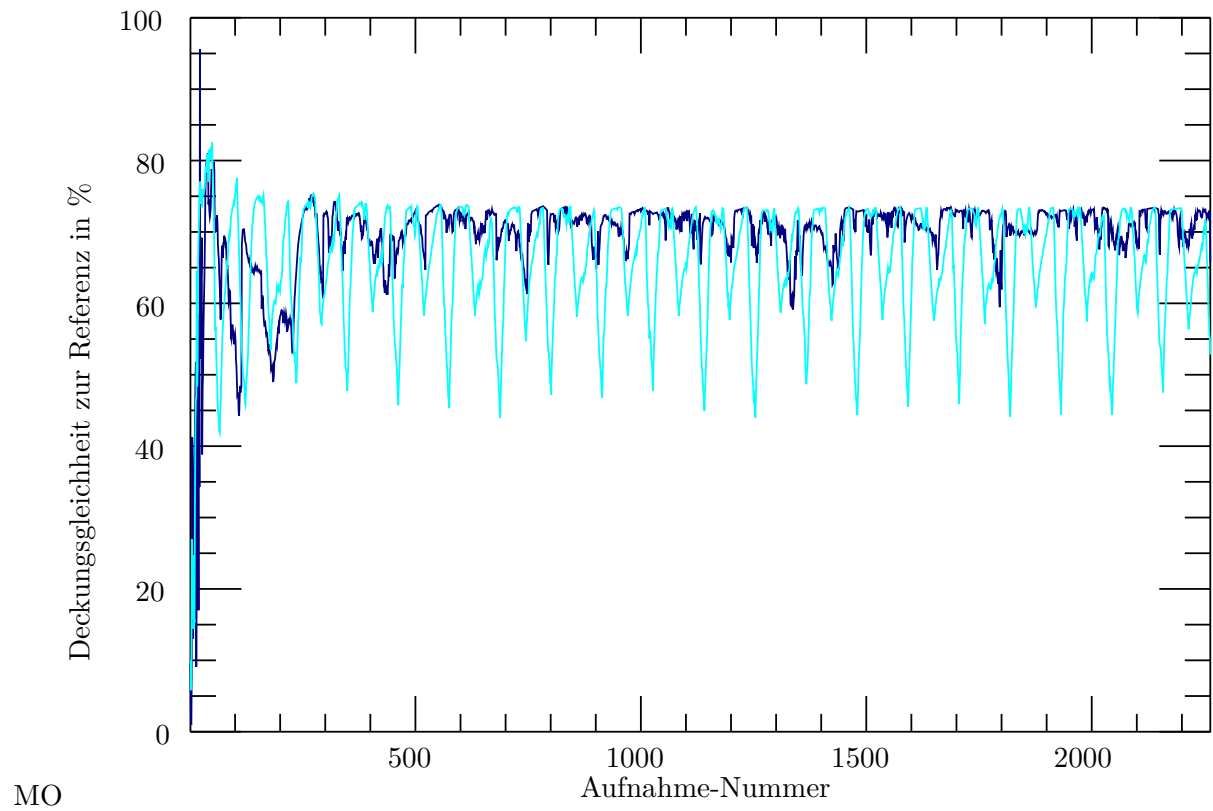


Abbildung 7.10: Verhältnis von Überlappungsvolumen V_U zum Referenzvolumen V_R für das Mittelpunkt-Tracking (hellblau) und Referenzpunkt-Tracking (dunkelblau) in einer simulierten partiell reduzierten Objektsicht und einem Messwerttrauschen von $\sigma = 0.2$ m bzw. rad (Szenario B)

7.3 Auflösung von Falsch- und Fehldetektion auf Objektebene

Für die in Abschnitt 6.2.3 vorgestellte Erkennung und Auflösung von Falsch- (FPO) und Fehldetektion (FNO) von Objekten konnte kein geeigneter Messaufbau gefunden werden, durch den eine statistisch belastbare Anzahl an Fehlern bestimmbar war. Vor allem der Einfluss weiterer Fehlerquellen, wie gegenseitige Verdeckungen von fahrenden Fahrzeugen, erschwert die nötige Fehlerisolation zur Untersuchung der einzelnen Messfehler.

Zur Veranschaulichung und Überprüfung des Verfahrens wurde daher eine Simulation aufgesetzt, die die Erkennung von Falsch- und Fehldetektion anhand unterschiedlicher Mess- und Sensorkonfigurationen verdeutlichen soll. Die Größe und Seitenverhältnisse des Messobjekts ähneln den fahrzeugtypischen Maßen ($4 \times 2 \times 2$ m). Das Messobjekt bewegt sich diskret entlang der Gitterpunkte eines Messfeldes analog zum Aufbau zur Genauigkeitsbestimmung der Sensoren (siehe Abschnitt 7.1 bzw. Abbildung 7.11).

Segmentierungswahrscheinlichkeit der Sensoren

Die Segmentierungswahrscheinlichkeiten (SW) der im FPO/FNO-Konflikt stehenden Sensoren werden, wie in Abschnitt 6.2.3 beschrieben, zur Bestimmung einer zusammengefassten Aussage herangezogen, um den Konflikt aufzulösen.

Im Zuge der Auswertung stellte sich hierbei eine Modellschwäche bei der Bestimmung der SW für das Stereokamerasystem heraus. Der in 4.1.3 beschriebene Ansatz führt beim Stereosystem zu einer Überschätzung der SW. Bei einer entsprechenden Kalibrierung des Modells durch eine Anpassung der Parameter für das Segmentierungsverhalten kommt es zu einer entsprechenden Verkleinerung der Segmentierungsräume (S-Ellipsoide). Dies führt wiederum zu einem nicht idealen Segmentierungsverhalten durch eine erhöhte Fragmentierung der Objektdaten. Dieser Offset zwischen dem Modell und der Realität lässt sich für das Stereoverfahren durch einige Faktoren erklären. Im Gegensatz zum Laserscanner liefert das verwendete Stereoverfahren keine dichte Punktmessung (z.B. Schwächen bei homogenen Flächen). Zudem nimmt die Stereodatenreduzierung Einfluss auf die Anzahl und Verteilung der Messpunkte. Im vorgestellten Modell zur Bestimmung der SW werden diese Faktoren nicht berücksichtigt und führen zu einer ungenauen SW-Modellierung beim Stereosystem.

Zur Kompensation der Modellschwäche wird ein Skalierungsfaktor s vorgeschlagen, der die modellbasiert bestimmbaren Abstände zwischen einem Objektpunkt P und seinen angrenzenden horizontalen und vertikalen Messpunkten \bar{P} und $\bar{\bar{P}}$ mit

$$\begin{aligned}\bar{P}_x &= \bar{P}_x \\ \bar{P}_y &= P_y + s(\bar{P}_y - P_y) \\ \bar{P}_z &= \bar{P}_z\end{aligned}\tag{7.1}$$

und

7.3 Auflösung von Falsch- und Fehldetektion auf Objektebene

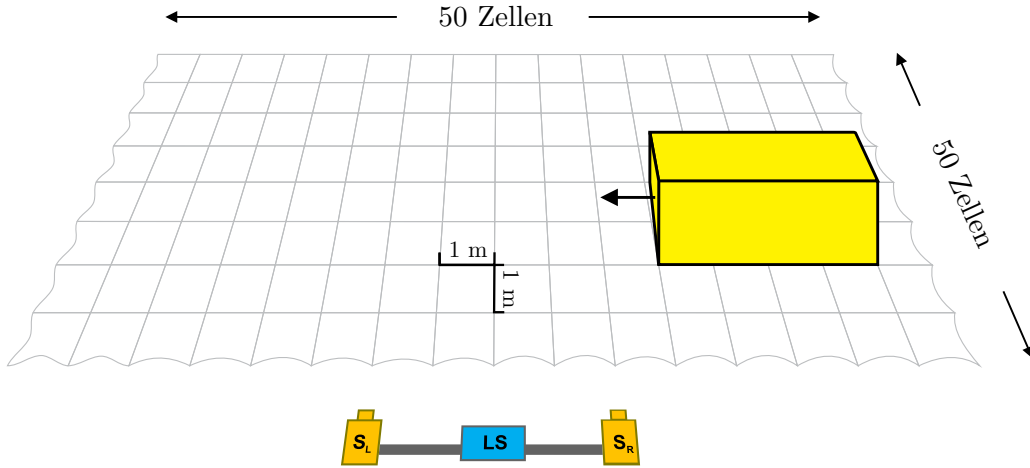


Abbildung 7.11: Darstellung des gitterartigen Messfeldes zur Simulation von Falsch- und Fehldetektion

$$\begin{aligned}
 \bar{\bar{P}}_x &= \bar{\bar{P}}_x \\
 \bar{\bar{P}}_y &= \bar{\bar{P}}_y \\
 \bar{\bar{P}}_z &= P_z + s(\bar{\bar{P}}_z - P_z)
 \end{aligned}
 \tag{7.2}$$

entsprechend skaliert (siehe hierzu Abschnitt 4.1.3).

Bei der Kalibrierung von s werden im ersten Schritt die Parameter zur Segmentierung entsprechend dem realen Messverhalten des Sensors angepasst. Anschließend wird s dahingehend bestimmt, dass sich ein, dem realen Messverhalten des Sensors, angepasstes Verhältnis der SW zwischen 0 % und 100 % ergibt. Für das eingesetzte Stereoverfahren wurde hierbei ein Faktor $s = 3.2$ verwendet. Das Segmentierungsmodell des Laserscanners bedurfte keiner Anpassung. Die Abbildung 7.12 stellt die entsprechend bestimmten SW der beiden Sensoren am jeweiligen Objektmittelpunkt im Messbereich dar (siehe Abbildung 7.11).

Die Segmentierungsunsicherheit im Nahbereich des Laserscanners resultiert aus der im gleichen Bereich erhöhten Messunsicherheit in Z-Richtung (siehe Abbildung 7.3(b)). Anschließend nehmen die SW für beide Sensoren mit zunehmender Messentfernung und steigendem Messrauschen ab.

Auswertung für zwei Sensoren

Ein FP-Objekt bzw. eine Falschdetektion eines Sensors wird bei einer FPO-Wahrscheinlichkeit von über 50 % angenommen (im Folgenden als rote Fläche dargestellt). Andernfalls wird angenommen, dass der Sensorkonflikt bei der Objektdetektion von einer Fehlmessung (FN-Objekt) verursacht wurde (grün dargestellte Fläche).

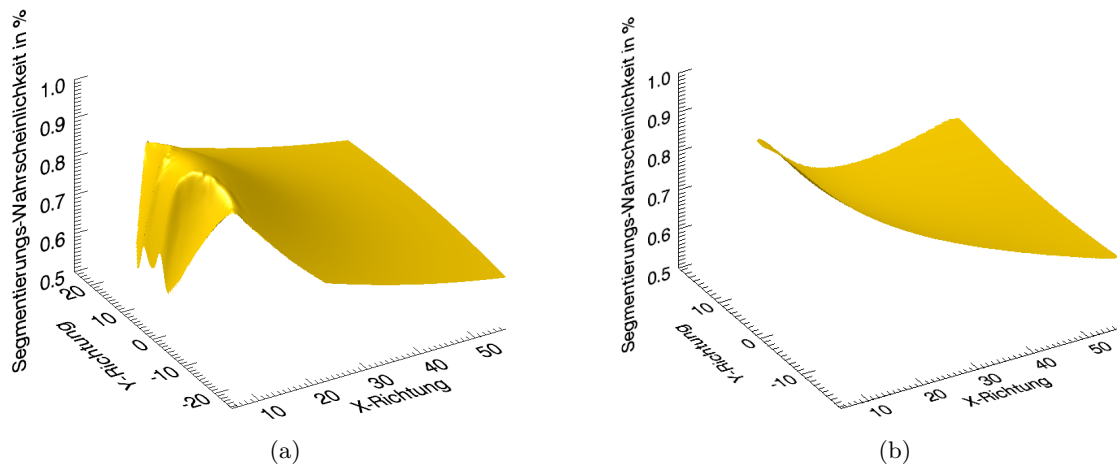


Abbildung 7.12: Segmentierungswahrscheinlichkeiten für den Laserscanner (a) und das Stereokamerasystem (b) für einen Messbereich von 51 m × 51 m in 1 m Höhe

Abbildung 7.13 illustriert die Wahrscheinlichkeit einer Falsch-Detektion des Messobjektes (FP-Objekt) für den Fall, dass entweder nur das Stereokamerasystem (a) oder nur der Laserscanner das Messobjekt detektiert hat (b). Abgesehen vom vorderen Messbereich bis 10 m, wo beide Sensoren sehr gegensätzliche SW besitzen, kann für den restlichen Messbereich keine signifikante allgemeingültige Konfliktursache bestimmt werden. Da die beiden Sensoren insgesamt ein ausgewogenes Segmentierungsverhalten besitzen, wird der Konflikt durch das lokale Verhältnis der beiden Sensor-SW aufgelöst. Hiermit bestätigt sich eine reelle Sensorkonkurrenz im Sinne von gleichberechtigten Sensoren.

Auswertung für drei Sensoren

Um jedoch eine signifikante Entscheidung zwischen einer Fehl- oder Falschdetektion treffen zu können, sollten mehr als zwei Sensoren eingebunden werden. In der Simulation ging hierfür die Detektionshypothese jeweils eines Sensors zweifach in die Auswertung ein, um einen dritten Sensor zu simulieren. Im Folgenden bezeichnet eine Detektionshypothese eines detektierenden Sensors eine positive Objektdetektion bzw. eine Nicht-Detektionshypothese eines nicht detektierenden Sensors eine negative Objektdetektion.

In Abbildung 7.14 sind die Ergebnisse mit zwei detektierenden Stereokamerasystemen und einem nicht detektierenden Laserscanner (a) sowie zwei detektierenden Laserscannern und einem nicht detektierenden Stereokamerasystem dargestellt. Abgesehen vom vorderen Messbereich bis 10 m bei der zweifachen Detektion durch den Laserscanner liegt die Wahrscheinlichkeit für eine korrekte Objektdetektion bzw. einer fehlenden Objektdetektion eines Sensors (FN-Objekt) für beide Konstellationen bei mindestens 60 %. Wie im vorderen Messbereich bei der zweiten Konstellation ersichtlich, ist es trotzdem noch möglich, dass eine sehr sichere Hypothese zwei unsichere Gegenhypothesen überstimmen kann (rot markierter Bereich).

7.3 Auflösung von Falsch- und Fehldetektion auf Objektebene

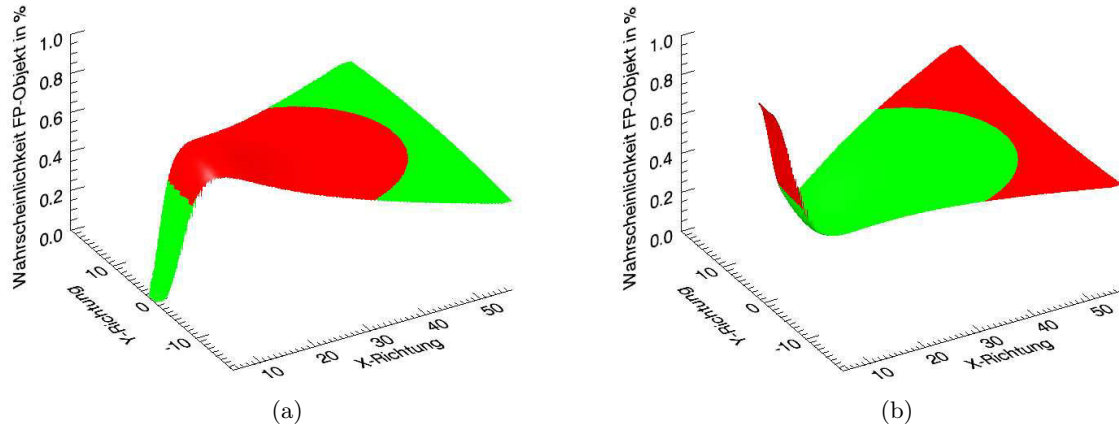


Abbildung 7.13: Wahrscheinlichkeiten für eine Objekt-Falschdetektion (FPO) bei einer alleinigen Objekterkennung durch das Stereokamerasystem (a) und durch den Laserscanner (b)

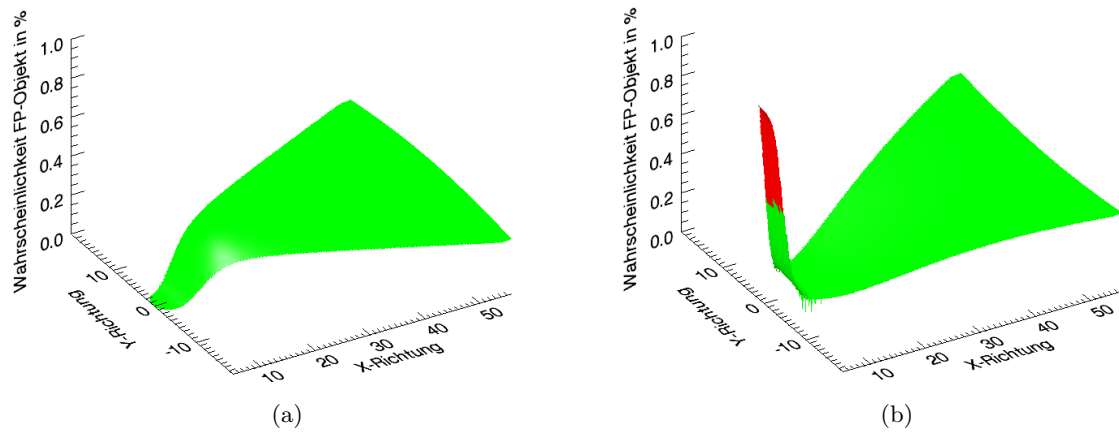


Abbildung 7.14: Wahrscheinlichkeiten für eine Objekt-Falschdetektion (FPO) bei einer Objekterkennung durch zwei Stereokamerasysteme gegenüber einer Nicht-Erkennung durch einen Laserscanner (a) sowie für eine Objekterkennung durch zwei Laserscanner gegenüber einer Nicht-Erkennung durch ein Stereokamerasystem (b)

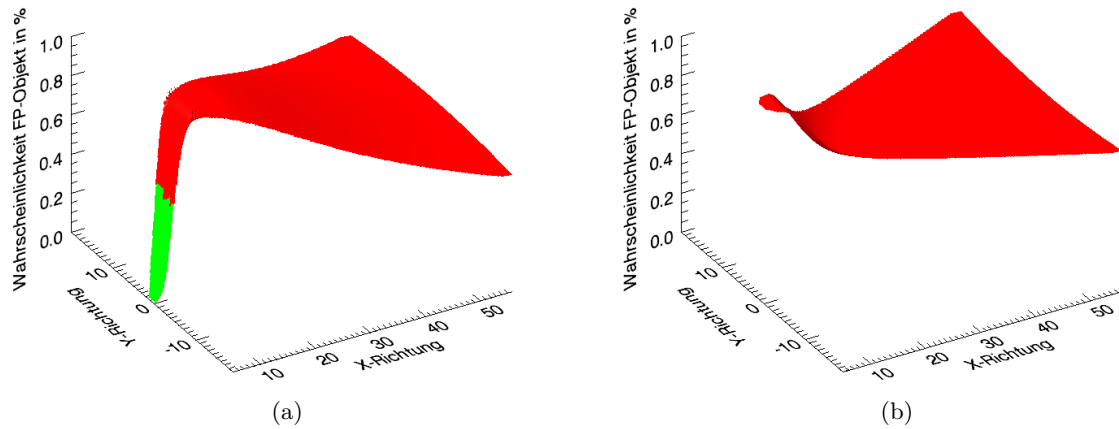


Abbildung 7.15: Wahrscheinlichkeiten für eine Objekt-Falschdetektion (FPO) bei Nicht-Erkennung durch zwei Laserscanner gegenüber einer Objekterkennung durch ein Stereokamerasystem (a) sowie für eine Nicht-Erkennung durch zwei Stereokamerasysteme gegenüber einer Objekterkennung durch einen Laserscanner (b)

Der gegensätzliche Konfliktfall von einer Detektionshypothese gegenüber zweier Nicht-Detektionshypothesen wird für die beiden Sensoren in Abbildung 7.15 dargestellt. Wie zu erwarten, treten hier zur Abbildung 7.14 an der 50 %-Ebene gespiegelte, identische Ergebnisse auf.

Des Weiteren kann davon ausgegangen werden, dass bei einer Vergrößerung des Verhältnisses zwischen Detektions- und Nicht-Detektionshypothesen eine noch höhere Eindeutigkeit in der Konfliktauflösung zu erwarten ist.

Für eine Auswertung zur Auflösung der Falsch- und Fehldetektion auf der Track-Ebene wurde wegen der Ähnlichkeit zum Verfahren auf der Objektebene verzichtet.

7.4 Auflösung der Fragmentierung und Zusammenfassung von Objekten

Da analog zur Erkennung von Fehl- und Falschdetektionen kein geeigneter Messaufbau gefunden werden konnte, um die Verfahren zur Erkennung und Bereinigung der Objektfragmentierung und -zusammenfassung mit realen Messdaten zu evaluieren, wurde eine entsprechende Simulation aufgesetzt.

Die Simulation wird wiederum auf dem in Abschnitt 7.1 vorgestellten Messfeld durchgeführt. Hierbei wird entweder ein Messobjekt ($4 \times 2 \times 2$ m) oder zwei gleich große Messobjekte ($2 \times 2 \times 2$ m) pro Messposition von den Sensoren detektiert. Wie in Abschnitt 6.2.2 beschrieben, wird modellhaft angenommen, dass bei der Betrachtung zweier Konflikt-Cluster C_1 und C_2 die Wahrscheinlichkeit für eine Objektfragmentierung fällt bzw. die Wahrscheinlichkeit für eine Objektzusammenfassung steigt, wenn der Abstand (d_{F-}) zwischen zwei Objekthypothesen eines Sensors in C_1 und C_2 steigt. Des Weiteren sinkt die Wahrscheinlichkeit für eine Objekt-

7.4 Auflösung der Fragmentierung und Zusammenfassung von Objekten

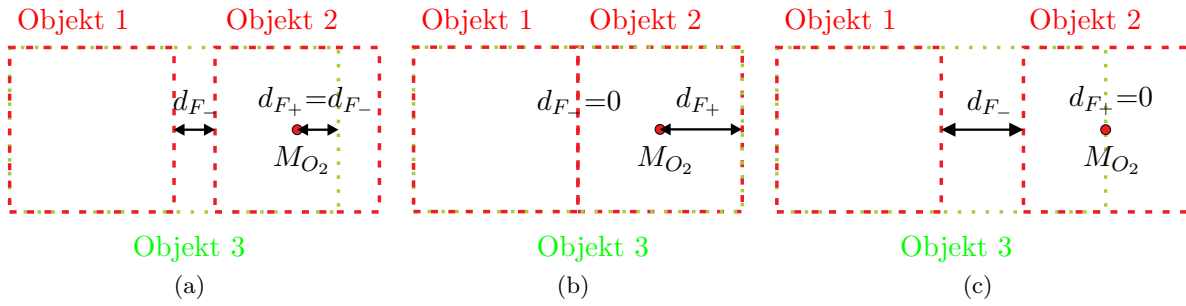


Abbildung 7.16: Simulation eines ausgeglichenen Verhältnisses der Wahrscheinlichkeiten zur Objektzusammenfassung und -fragmentierung ((a) – Szenario A), einer hohen Wahrscheinlichkeit zur Objektfragmentierung ((b) – Szenario B) und einer hohen Wahrscheinlichkeit zur Objektzusammenfassung ((c) – Szenario C) über die Variation der Trennungs- und Überlappungsdistanz d_{F-} und d_{F+} dreier Objekthypothesen

zusammenfassung bzw. steigt die Wahrscheinlichkeit für eine Objektfragmentierung innerhalb C_1 und C_2 mit einer steigenden Überlappung einer Objekthypothese eines Sensors über C_1 und C_2 . Der Grad der Überlappung bestimmt sich aus dem kürzesten Abstand (d_{F+}) des Cluster-Mittelpunkts (M_{O_2}) zum Rand des überlappenden Objekts.

Es wurden drei verschiedene Messszenarien aufgestellt. **Szenario A** stellt eine gleichgewichtete Wahrscheinlichkeit für eine Objektzusammenfassung und Objektfragmentierung (siehe Abbildung 7.16(a)) dar. Die Trennungsdistanz d_{F-} zwischen den Objekthypothesen 1 und 2 des ersten Sensors entspricht hierbei der Überlappungsdistanz d_{F+} der Objekthypothese 3 des zweiten Sensors. **Szenario B** simuliert eine höhere Wahrscheinlichkeit für eine Objektfragmentierung (siehe Abbildung 7.16(b)) und **Szenario C** beschreibt eine höhere Wahrscheinlichkeit für eine Objektzusammenfassung (siehe Abbildung 7.16(c)) über eine Ungleichgewichtung der entsprechend simulierten Überlappungs- und Trennungsdistanzen.

Auswertung für zwei Sensoren

Bei einer simulierten Messung mit einem Laserscanner und einem Stereokamerasystem und deren Modellparametern ergibt sich für das ausgeglichene Szenario A (siehe Abbildung 7.16(a)) die in Abbildung 7.17 dargestellte Wahrscheinlichkeit für ein Zusammenfassungs-Objekt (ZO) zwischen zwei Konflikt-Clustern. Hierbei überwiegt bei beiden Messszenarien bis zu einer Messtiefe von ca. 30 m die Wahrscheinlichkeit für ein ZO. Dies resultiert aus der für beide Sensoren in diesem Bereich hohen SW, wodurch sich die Wahrscheinlichkeit für eine Objektzusammenfassung erhöht.

Da für die gewählten Messobjektgrößen die Distanzen d_{F-} und d_{F+} bei 0.5 m liegen und die Fehlerellipsen für beide Sensoren um mindestens eine Größenordnung kleiner sind, gehen die daraus resultierenden Trennungs- und Überlappungswahrscheinlichkeiten $P^-(F_-)$ und $P^-(F_+)$ gegen 1. Somit bestimmt beim Szenario eines ausgeglichenen FO-ZO-Verhältnisses nur die SW die resultierende Gesamtwahrscheinlichkeit.

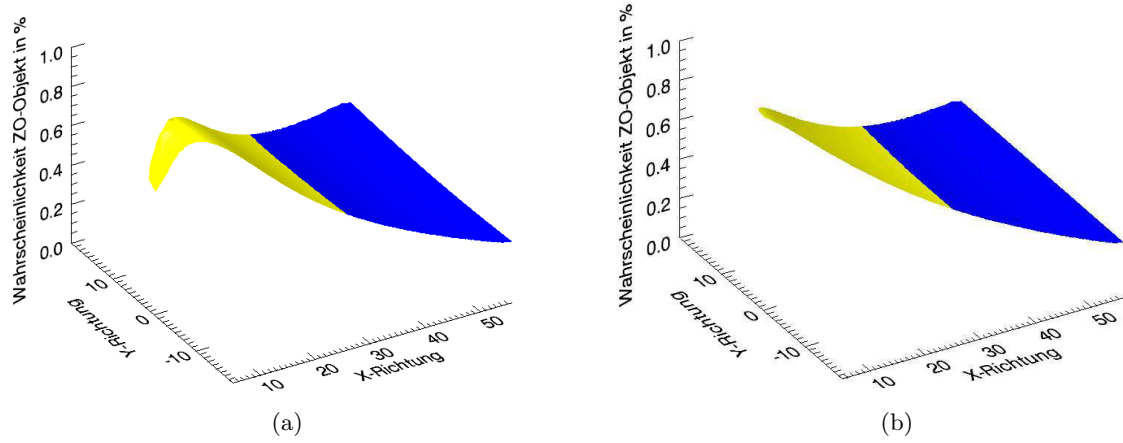


Abbildung 7.17: Wahrscheinlichkeiten für ein Zusammenfassungs-Objekt (ZO) bei einer Erkennung einer Objektzusammenfassung vom Stereokamerasystem gegenüber einer Erkennung einer Objektfragmentierung mit gleich hoher Wahrscheinlichkeit vom Laserscanner (a) und umgekehrt (b) entsprechend dem Szenario A

Verringern sich die Werte von d_{F-} oder d_{F+} , wie in den nicht ausgeglichenen FO-ZO-Simulationen (siehe Abbildungen 7.16(b) und 7.16(c)), so wirkt sich dies in einer stärkeren Ausprägung der Gesamtwahrscheinlichkeit in eine entsprechende Richtung aus.

In Abbildung 7.18 wird dies für die Simulation einer erhöhten Wahrscheinlichkeit zur Objektfragmentierung durch eine im Vergleich zu d_{F-} größeren Distanz d_{F+} ersichtlich. Sowohl für den Laserscanner (a) als auch das Stereokamerasystem als fragmentierenden Sensor (b) deutet die Gesamtwahrscheinlichkeit klar auf ein FO hin.

Die Simulation einer erhöhten Wahrscheinlichkeit zur Objektzusammenfassung durch eine im Vergleich zu d_{F+} größeren Distanz d_{F-} resultiert erwartungsgemäß ebenso in einer erhöhten Gesamtwahrscheinlichkeit für ein ZO (siehe Abbildung 7.19).

7.4 Auflösung der Fragmentierung und Zusammenfassung von Objekten

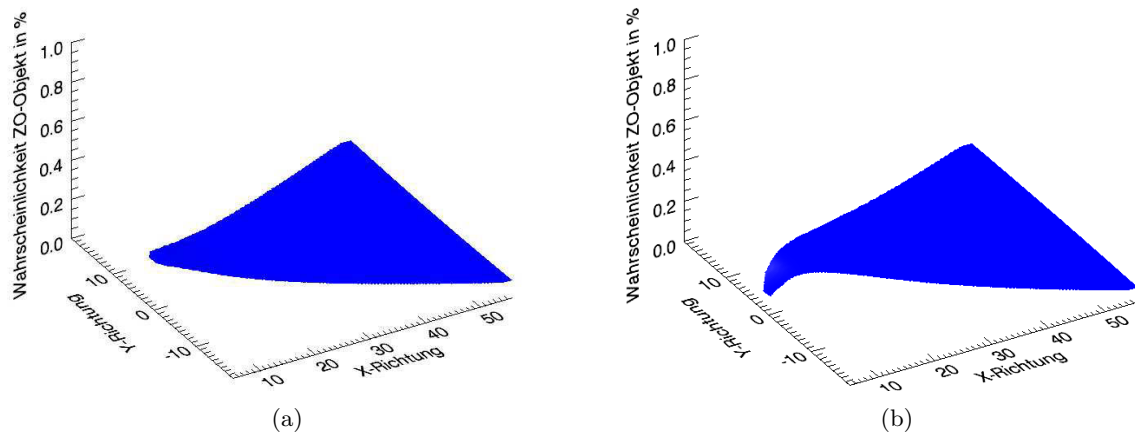


Abbildung 7.18: Wahrscheinlichkeiten für ein Zusammenfassungs-Objekt (ZO) bei einer Erkennung einer Objektfragmentierung durch ein Stereokamerasystem gegenüber einer Erkennung einer Objektzusammenfassung mit höherer Wahrscheinlichkeit durch einen Laserscanner (a) sowie bei einer Erkennung einer Objektfragmentierung durch einen Laserscanner gegenüber einer Erkennung einer Objektzusammenfassung mit höherer Wahrscheinlichkeit durch ein Stereokamerasystem (b) entsprechend dem Szenario B

7 Experimentelle Ergebnisse

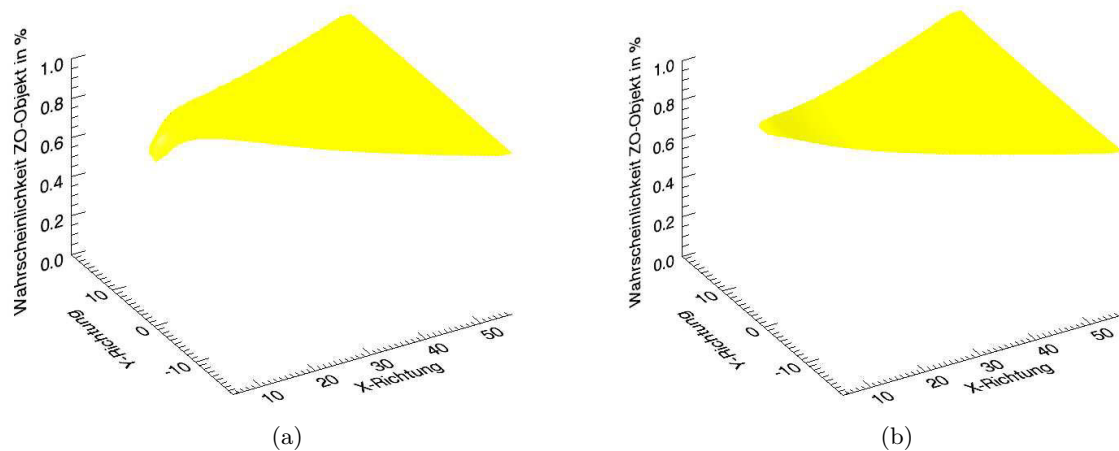


Abbildung 7.19: Wahrscheinlichkeiten für ein Zusammenfassungs-Objekt (ZO) bei einer Erkennung einer Objektfragmentierung durch ein Stereokamerasystem gegenüber einer Erkennung einer Objektzusammenfassung mit geringerer Wahrscheinlichkeit durch einen Laserscanner (a) sowie bei einer Erkennung einer Objektfragmentierung durch einen Laserscanner gegenüber einer Erkennung einer Objektzusammenfassung mit geringerer Wahrscheinlichkeit durch ein Stereokamerasystem (b) entsprechend dem Szenario C

Auswertung für drei Sensoren

Um zu zeigen, dass sich für mehr als zwei eingehende Sensoren die Unsicherheit der Gesamtwahrscheinlichkeit bzw. der Konfliktauflösung verringert, wurde für die Simulation des ausgewogenen FO-ZO-Verhältnisses (Szenario A) die Detektionshypothese eines Sensors wiederum doppelt gewertet. Die Ergebnisse für eine zweifache Einzelobjektdetektion im Gegensatz zu einer Doppelobjektdetektion sowie eine zweifache Doppelobjektdetektion im Gegensatz zu einer Einzelobjektdetektion sind in den Abbildungen 7.20 und 7.21 dargestellt. In beiden Konstellationen zeigen sich eindeutige Trendverschiebungen für eine FO- bzw. ZO-Detektion, die durch die entsprechend doppelt eingehenden Detektions-Hypothesen hervorgerufen werden.

Auf die Darstellung der Simulationsergebnisse von Szenario B und C mit mehr als zwei eingehenden Sensoren wurde verzichtet, da sich die eindeutigen Tendenzen aus der zweifachen Sensorauswertung hier konsequenterweise weiter verstärken.

7.4 Auflösung der Fragmentierung und Zusammenfassung von Objekten

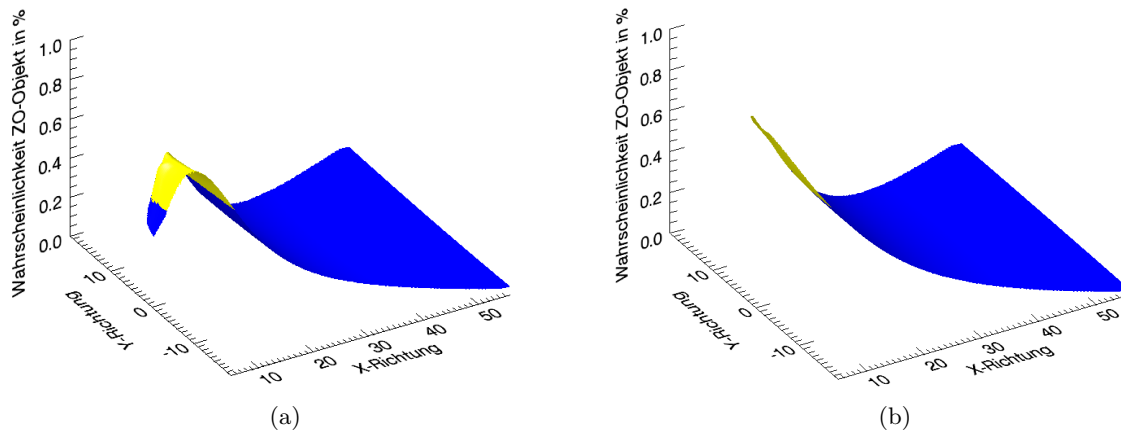


Abbildung 7.20: Wahrscheinlichkeiten für ein Zusammenfassungs-Objekt (ZO) bei einer Erkennung einer Objektzusammenfassung durch zwei Stereokamerasysteme gegenüber einer Erkennung einer Objektfragmentierung durch einen Laserscanner (a) sowie bei einer Erkennung einer Objektzusammenfassung durch zwei Laserscanner gegenüber einer Erkennung einer Objektfragmentierung durch ein Stereokamerasystem (b) für das Szenario A

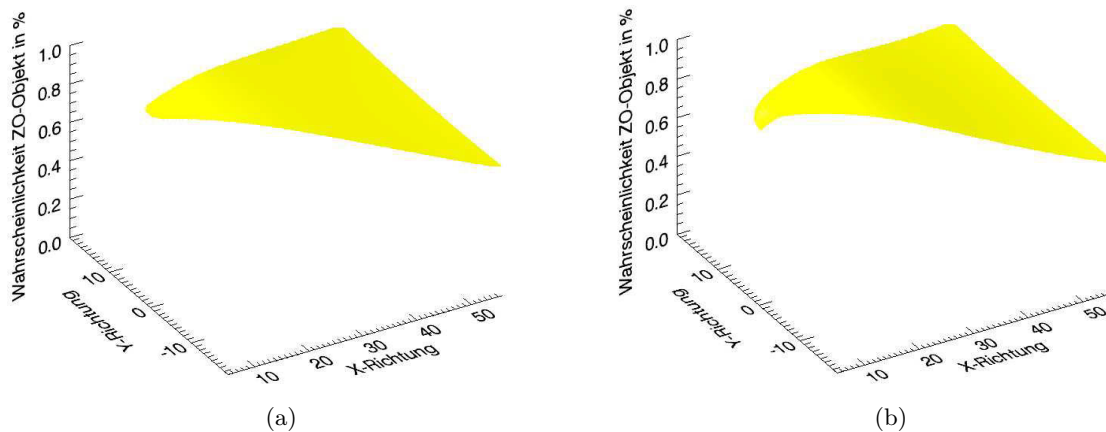


Abbildung 7.21: Wahrscheinlichkeiten für ein Zusammenfassungs-Objekt (ZO) bei einer Erkennung einer Objektfragmentierung durch zwei Stereokamerasysteme gegenüber einer Erkennung einer Objektzusammenfassung durch einen Laserscanner (a) sowie bei einer Erkennung einer Objektfragmentierung durch zwei Laserscanner gegenüber einer Erkennung einer Objektzusammenfassung durch ein Stereokamerasystem (b) für das Szenario A

7.5 Auflösung von Multi-Objekt-Tracks

Das in Abschnitt 6.3.3 vorgestellte Verfahren zur Erkennung von Multi-Objekt-Tracks (MOT) wurde ebenfalls in zwei simulierten Szenarien untersucht.

In **Szenario 1** (siehe Abbildung 7.22) kreuzen sich zwei mit 50 km/h linear bewegende Objekte in einem Winkel von 45° in der Mitte ihrer Wege. Ein korrekt assoziierender Tracker eines Sensors S_1 verwendet als Messwerte für jeweils einen Track die Objektdetektionen eines Objektes (Einzel-Objekt-Track – EOT). Ein fehlerhaft assoziierender Tracker eines Sensors S_2 verfolgt bis zum Kreuzungspunkt die Messwerte des einen Objekts und anschließend die Messwerte des zweiten Objekts (Multi-Objekt-Track – MOT). Jeweils ein Sensor detektiert somit entweder zwei EOT oder zwei MOT über den gesamten Messverlauf.

Das **Szenario 2** (siehe Abbildung 7.23) simuliert eine alternative Situation für eine Falsch-zuordnung während einer Assoziation. Hierbei fahren zwei Objekte im Abstand von 5 m mit einer Geschwindigkeit von 50 km/h parallel zueinander in einer linearen Bewegung. Am Kreuzungspunkt ordnen die fehlerhaft assoziierenden Tracker des einen Sensors die beiden Objekte den falschen Tracks zu, wohingegen die Tracker des anderen Sensors die Objekte den korrekten Tracks weiter zuordnen.

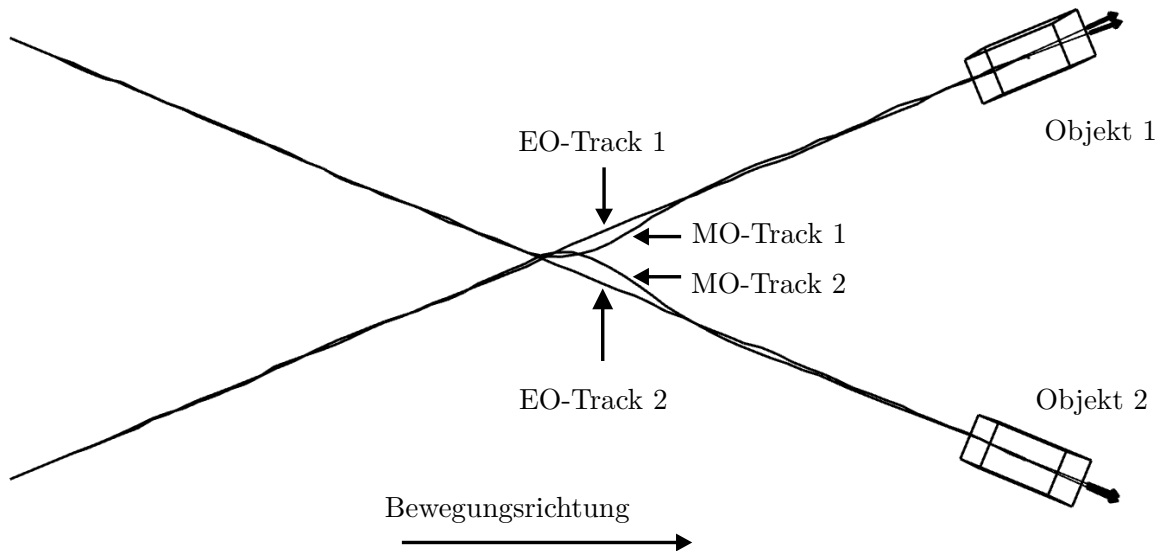


Abbildung 7.22: Szenario 1: Zwei Objekte kreuzen sich in einem Winkel von 45° mit 50 km/h, wobei Sensor 1 zwei fehlerfreie EOT und Sensor 2 zwei fehlerhafte MOT detektiert (Abbildung mit simulierter Messgenauigkeit von $\sigma = 0.1$ m)

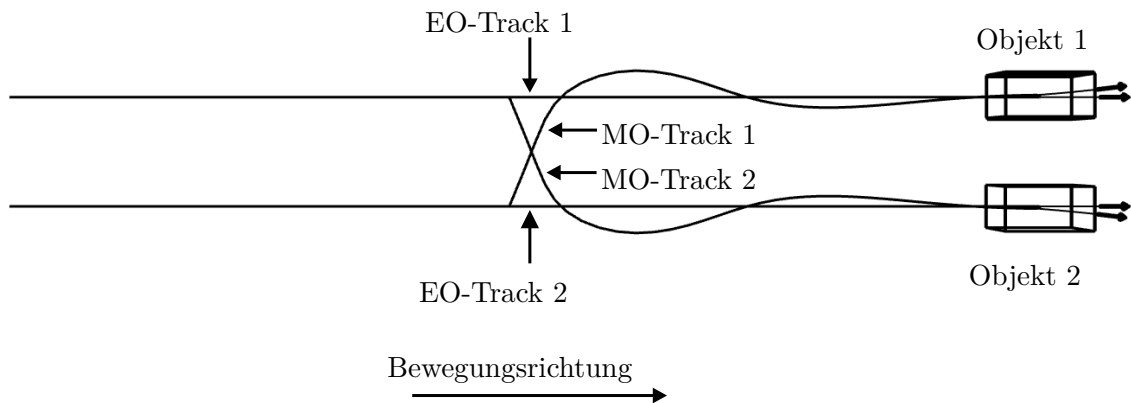


Abbildung 7.23: Szenario 2: Zwei Objekte fahren im Abstand von 5 m parallel zueinander mit 50 km/h, wobei Sensor 1 zwei fehlerfreie EOT und Sensor 2 zwei fehlerhafte MOT detektiert (Abbildung mit simulierter Messgenauigkeit von $\sigma = 0.1$ m)

Auswertung für zwei Sensoren

Die beiden Szenarien wurden zunächst mit zwei Sensoren bzw. jeweils zwei Trackern untersucht. Das Ziel der Simulation war die Bestimmung des MOT-verursachenden Sensors durch den ermittelten System-Cluster. Hierzu wurden die Messwerte der Objektpositionen zusätzlich mit einer normalverteilten Zufallsgröße verrauscht, um die Abhängigkeit der MOT-Bestimmung zur Messgenauigkeit des Sensors zu überprüfen.

In Tabelle 7.3 sind die simulierten Ergebnisse für einen MOT- und einen EOT-detektierenden Sensor in Szenario 1 dargestellt. Die prozentuale Angabe entspricht dem Anteil der korrekt erkannten MO-Tracks bezogen auf die gesamte Anzahl der Simulationsdurchläufe. Für jede Konstellation wurden 1000 Durchläufe simuliert. Bei einer gering simulierten Streuung der Objektposition von $\sigma = 0.01$ m für beide Sensoren liegt die MOT-Erkennung bei 100 %. Diese erstaunlich hohe Quote resultiert aus dem Verhalten des MOT im Vergleich zum EOT. Durch die schlagartige Änderung der Bewegungsrichtung und der Abweichung vom aktuellen Track-Zustand und der Track-Vorhersage benötigt der Tracker eine gewisse Laufzeit, um sich auf die neue Bewegung einzustellen. Dieses vom Idealverlauf abweichende Verhalten wirkt sich nachteilig für den MOT in der Assoziierung vom Track-Cluster zum System-Cluster aus. Die beiden EO-Tracks werden somit auch nach dem MOT-Effekt weiterhin dem ursprünglichen System-Cluster zugeordnet, wohingegen die beiden MO-Tracks dem jeweils anderen System-Cluster zugeordnet werden, wodurch sie günstiger Weise als fehlerhafte Tracks erkannt werden.

Bei einer gleich starken Streuung der Objektposition für beide Sensoren nimmt die MOT-Erkennung mit zunehmender Varianz leicht ab (siehe Hauptdiagonale in Tabelle 7.3). Bei einer stärkeren Streuung der Objektposition für den EOT-Sensor werden die EO-Tracks in der System-Cluster-Zuordnung aufgrund ihrer höheren Varianzen benachteiligt und fälschlicherweise als MO-Tracks erkannt, was die Fähigkeit zur ambivalenten Konfliktauflösung des Verfahrens verdeutlicht.

	EOT $\sigma = 0.01$ m	EOT $\sigma = 0.1$ m	EOT $\sigma = 0.2$ m
MOT $\sigma = 0.01$ m	100 %	0 %	0 %
MOT $\sigma = 0.1$ m	100 %	97.6 %	0 %
MOT $\sigma = 0.2$ m	100 %	100 %	89.3 %

Tabelle 7.3: Simulation der Erkennung eines MOT-Effekts für zwei gegensätzlich trackende Sensoren mit unterschiedlichen Messgenauigkeiten (Szenario 1)

Die gleichen Aussagen lassen sich auch für die Simulation in Szenario 2 ableiten (siehe Tabelle 7.4). Im Gegensatz zu Szenario 1 liegt die MOT-Erkennungsquote bei gleich starker Streuung der Objektpositionen unabhängig von der Streuungsstärke bei 100 % (siehe Hauptdiagonale in Tabelle 7.4). Dies lässt sich durch die stärkere Abweichung des MOT vom idealen Track nach dem Objektwechsel erklären, welche die System-Cluster-Bildung auch bei einer höheren Messungenauigkeit stärker als in Szenario 1 beeinflusst.

	EOT $\sigma = 0.01$ m	EOT $\sigma = 0.1$ m	EOT $\sigma = 0.2$ m
MOT $\sigma = 0.01$ m	100 %	0 %	0 %
MOT $\sigma = 0.1$ m	100 %	100 %	0 %
MOT $\sigma = 0.2$ m	100 %	100 %	100 %

Tabelle 7.4: Simulation der Erkennung eines MOT-Effekts für zwei gegensätzlich trackende Sensoren mit unterschiedlichen Messgenauigkeiten (Szenario 2)

Auswertung für drei Sensoren

Neben der Genauigkeit der Tracks sollte auch das mengenmäßige Verhältnis der beiden widersprüchlichen Track-Mengen bzw. der eingehenden Sensoren, die in einem MOT-Konflikt stehen, einen Einfluss auf die MOT-Bewertung haben. So sollte bei gleich genauen Tracks, die größere Menge der beiden im Konflikt stehenden Track-Mengen in ihrer Gesamtaussage höher bewertet werden. Um dies zu überprüfen, wurden für die beiden simulierten Szenarien entweder die EO-Tracks oder die MO-Tracks zweifach in die Auswertung einbezogen.

Die Tabellen 7.5 und 7.6 stellen hierbei die Simulationsergebnisse für zweifach gewertete EOT in Szenario 1 und 2 dar. Bei gleicher oder höherer Genauigkeit der EOT wird eine 100-prozentige korrekte Erkennung der fehlerhaften MOT erzielt. Ist der MOT genauer, so werden die anderen beiden EO-Tracks fälschlicherweise als MOT gewertet. Wie in Abschnitt 6.3.3 beschrieben, kann der Einfluss der Track-Genauigkeit auf die MOT-Erkennung über den Gewichtungsfaktor f (siehe Gleichung 6.26) vermindert und der Einfluss des mengenmäßigen Verhältnisses zwischen den Konflikt-Track-Mengen erhöht werden. In der Simulation wurde $f = 1$ gesetzt und damit keine Gewichtung vorgenommen.

	EOT $\sigma = 0.01$ m	EOT $\sigma = 0.1$ m	EOT $\sigma = 0.2$ m
MOT $\sigma = 0.01$ m	100 %	0 %	0 %
MOT $\sigma = 0.1$ m	100 %	100 %	0 %
MOT $\sigma = 0.2$ m	100 %	100 %	100 %

Tabelle 7.5: Simulation der Erkennung eines MOT-Effekts für einen MOT-detektierenden und zwei EOT-detektierende Sensoren mit unterschiedlichen Messgenauigkeiten (Szenario 1)

	EOT $\sigma = 0.01$ m	EOT $\sigma = 0.1$ m	EOT $\sigma = 0.2$ m
MOT $\sigma = 0.01$ m	100 %	0 %	0 %
MOT $\sigma = 0.1$ m	100 %	100 %	0 %
MOT $\sigma = 0.2$ m	100 %	100 %	100 %

Tabelle 7.6: Simulation der Erkennung eines MOT-Effekts für einen MOT-detektierenden und zwei EOT-detektierende Sensoren mit unterschiedlichen Messgenauigkeiten (Szenario 2)

7 Experimentelle Ergebnisse

Bei einer zweifachen Wertung der MOT ergibt sich ein komplementäres Resultat (siehe Tabellen 7.7 und 7.8). Im Falle von vergleichsweise genaueren oder gleich genauen MO-Tracks werden die EO-Tracks fälschlicherweise als MOT bewertet. Nur bei einer gemeinsamen Messwertstreuung von $\sigma = 0.2$ m fällt diese Entscheidung knapp aus, da hier wiederum bedingt durch das Szenario und der Verzerrung der Track-Entwicklung bei den MOT, die EOT in der System-Cluster-Bestimmung begünstigt werden.

Somit wird bei gleicher oder höherer Genauigkeit von zwei Tracks mit gleichem Track-Verlauf T_1 im Vergleich zu einem Track mit einem alternativen Track-Verlauf T_2 in einem gemeinsamen System-Cluster der Verlauf T_1 als richtig (EOT) und T_2 als fehlerhaft (MOT) bestimmt.

	EOT $\sigma = 0.01$ m	EOT $\sigma = 0.1$ m	EOT $\sigma = 0.2$ m
MOT $\sigma = 0.01$ m	0 %	0 %	0 %
MOT $\sigma = 0.1$ m	100 %	0 %	0 %
MOT $\sigma = 0.2$ m	100 %	100 %	45.1 %

Tabelle 7.7: Simulation der Erkennung eines MOT-Effekts für einen EOT-detektierenden und zwei MOT-detektierende Sensoren mit unterschiedlichen Messgenauigkeiten (Szenario 1)

	EOT $\sigma = 0.01$ m	EOT $\sigma = 0.1$ m	EOT $\sigma = 0.2$ m
MOT $\sigma = 0.01$ m	0 %	0 %	0 %
MOT $\sigma = 0.1$ m	100 %	0 %	0 %
MOT $\sigma = 0.2$ m	100 %	100 %	0 %

Tabelle 7.8: Simulation der Erkennung eines MOT-Effekts für einen EOT-detektierenden und zwei MOT-detektierende Sensoren mit unterschiedlichen Messgenauigkeiten (Szenario 2)

7.6 Genauigkeit der Objektbestimmung

Zur Betrachtung der Objektbestimmungsgüte wird zum einen die Ungenauigkeit in der Objektbestimmung (OBF) auf der mittleren Prozessebene sowie die Ungenauigkeit in der Objektsstandsbestimmung (OZF) auf der oberen Prozessebene untersucht.

Um die Genauigkeit der Objektbestimmung in Abhängigkeit der eingesetzten Fusionsebene unter realen Bedingungen zu ermitteln, wurde ein zweites Fahrzeug als Referenzfahrzeug mit zwei hochgenauen GPS-Empfängern ausgestattet (*Trimble 5700*). Die beiden RTK-fähigen (*Real Time Kinematic*) GPS-Empfänger zeichnen beim gleichzeitigen Empfang von Basisinformationen ihre Position mit einer Frequenz von 10 Hz und einer Positionsgenauigkeit von unter 2 cm (σ_1) auf [107]. Die Position und Ausrichtung der beiden Antennen am Fahrzeug, als auch das Fahrzeug selbst, wurden im Vorfeld genau vermessen (Ausmaße des Referenzfahrzeugs siehe Tabelle 7.9). Somit ist es möglich unter der Annahme einer ebenen Fahrbahn die Fahrzeugform und

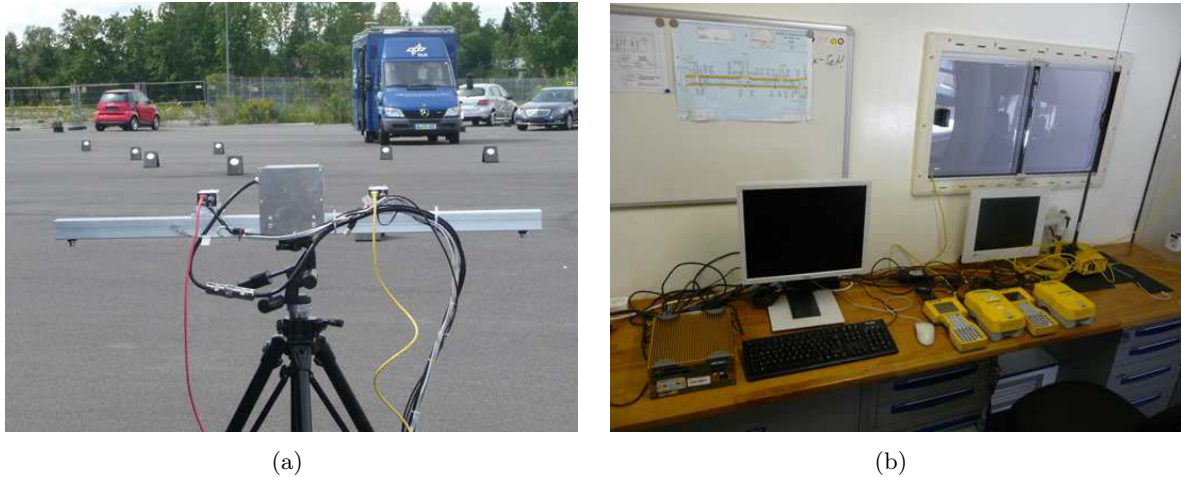


Abbildung 7.24: Aufbau der Messmarken zur Bestimmung der Transformation von SKS in UTM-KS (a); Messaufbau im Inneren des Referenzfahrzeugs (b)

somit die umschließende, ideale Objektbox des Referenzfahrzeugs aus den beiden Antennenpositionen abzuleiten. Um weitere Fehlereinflüsse zu vermeiden, blieb das eigentliche Messfahrzeug während der Messung stehen und zeichnete die Bewegungen des vor ihm fahrenden Referenzfahrzeugs auf.

Zum Vergleich der aus den Sensordaten und der Datenfusion abgeleiteten Objekte mit den Referenzobjekten des Referenzfahrzeugs, müssen die Referenzdaten in das Sensorkoordinatensystem SKS, welches durch die linke Kamera definiert ist, transformiert werden. Hierfür wurde im Vorfeld der Messung die äußere Orientierung der linken Kamera im Bezug zum UTM-Koordinatensystem bestimmt. Für diesen Zweck wurden im Sichtbereich der Kamera Messmarken platziert und verortet und die äußere Orientierung über einen Rückwärtsschnitt bestimmt (siehe Abbildung 7.24). Die Fehler in der Verortung der Messmarken lag hierbei im Mittel unter 1 cm.

Als Maßstab der Bestimmungsgüte eines Objektes O_M wurde das Überdeckungsvolumen von O_M zu seinem Referenzobjekt gewählt. Somit fließt die Güte der einzelnen Zustandsparameter, wie die Objektausmaße, -position und -ausrichtung in die Betrachtung mit ein. Für alle Tracking-Filter wurden hierbei identische Parametersätze verwendet (siehe Anhang).

d_x	d_y	d_z	V
6.41 m	2.202 m	2.69 m	37.97 m^3

Tabelle 7.9: Ausmaße und Volumen des Referenzfahrzeugs

Insgesamt wurden drei verschiedene Messfahrten mit unterschiedlichen Routencharakteristika durchgeführt. Die Durchschnittsgeschwindigkeit lag bei ca. 10 km/h.

Route 1 (siehe Abbildung 1 im Anhang) stellt eine geschlossene Dreiecksfahrt des Referenzfahrzeugs mit 13 Wiederholungen dar und besitzt einen entsprechend hohen Anteil an linearen Bewegungen.

Route 2 entspricht einer geschlossenen Kreisfahrt in 11-facher Wiederholung, um eine konstante nichtlineare Bewegung des Referenzfahrzeugs herbeizuführen (siehe Abbildung 2 im Anhang).

Route 3 beschreibt eine kurvige Fahrt mit einer hohen nichtlinearen Bewegungsdynamik (siehe Abbildung 3 im Anhang). Route 1 wurde mit 6334 Einzelaufnahmen, Route 2 mit 4063 Einzelaufnahmen und Route 3 mit 2119 Einzelaufnahmen durch die beiden Sensoren vermessen.

7.6.1 Vergleich der Sensoren

Beim Vergleich der unfusionierten Sensordaten ergibt sich eine insgesamt bessere Objektschätzung durch das Stereosystem (siehe Tabelle 7.10), was sich vor allem durch den schmalen Messbereich und der groben Auflösung des Laserscanners in Z-Richtung erklären lässt.

Im Vergleich zu den ungetrackten Sensordaten (siehe Abbildungen 7.25, 7.26 und 7.27) steigt die Güte der Objektschätzung für den Laserscanner durch das Tracking (siehe Abbildungen 7.28, 7.29 und 7.30). Beim Stereosystem ist eine Genauigkeitssteigerung durch das Tracking nur bei Route 1 zu erkennen. Die schlechtere Objektschätzung der getrackten Stereodaten in Fahrt 3 lässt sich durch eine verzögerte Adaption des getrackten Objektzustands an die hohe Bewegungsdynamik des Referenzfahrzeugs erklären.

Die zu den anderen Routen vergleichsweise ungenauere Objektschätzung durch die Sensoren in Route 1 begründet sich durch die routenbedingte schlechte Beobachtbarkeit des Referenzfahrzeugs auf den geraden Streckenabschnitten.

	Stereo ohne Tracking	Stereo mit Tracking	Laser ohne Tracking	Laser mit Tracking
Fahrt 1	22.33 m^3	22.52 m^3	29.31 m^3	28 m^3
Fahrt 2	18.45 m^3	17.59 m^3	26.04 m^3	24.68 m^3
Fahrt 3	17.92 m^3	20.57 m^3	25.20 m^3	24.65 m^3

Tabelle 7.10: Mittlere Abweichung der Objektbox-Volumina zwischen den ungetrackten Messobjekten der Sensoren und den Referenzobjekten

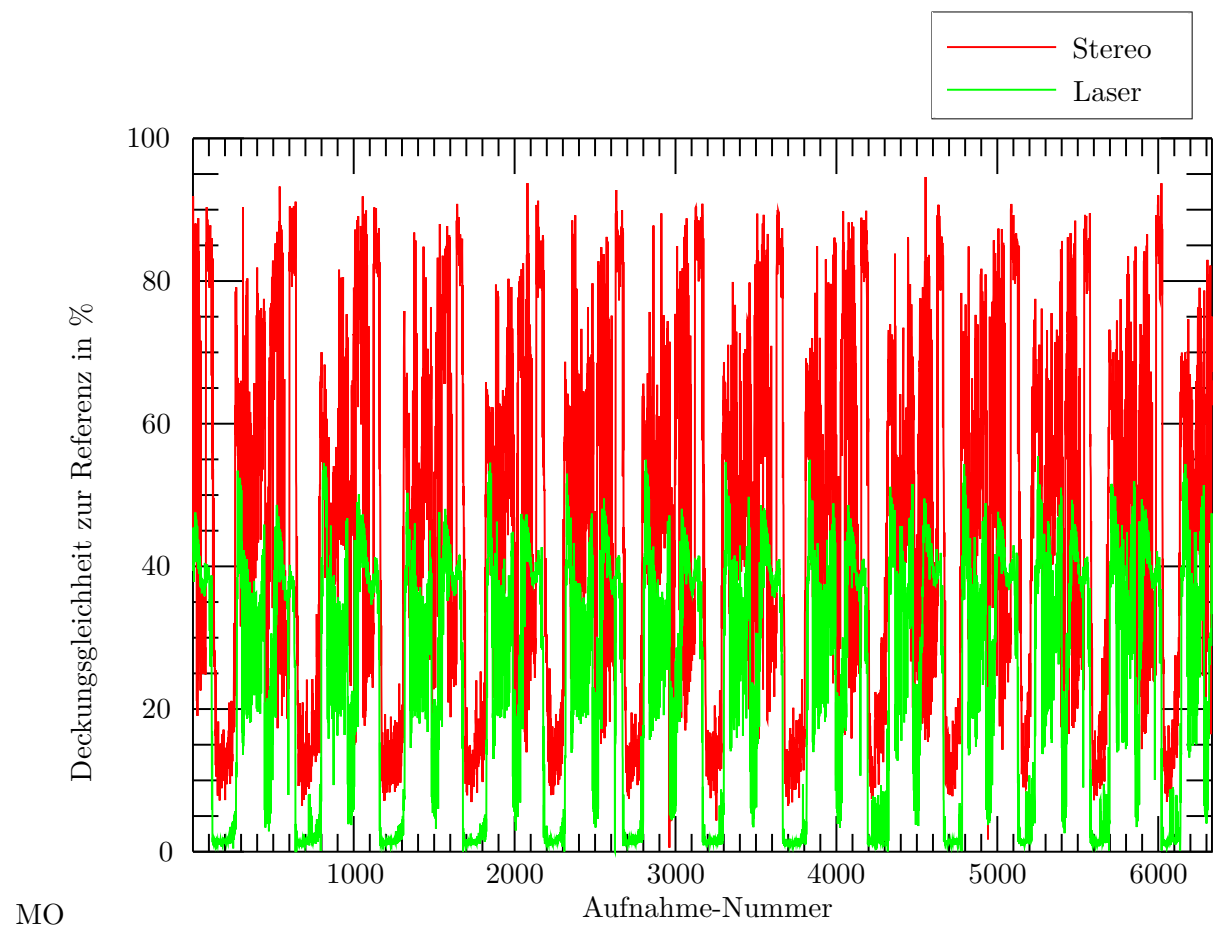


Abbildung 7.25: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den ungetrackten Messobjekten der Sensoren und den Referenzobjekten für die Route 1

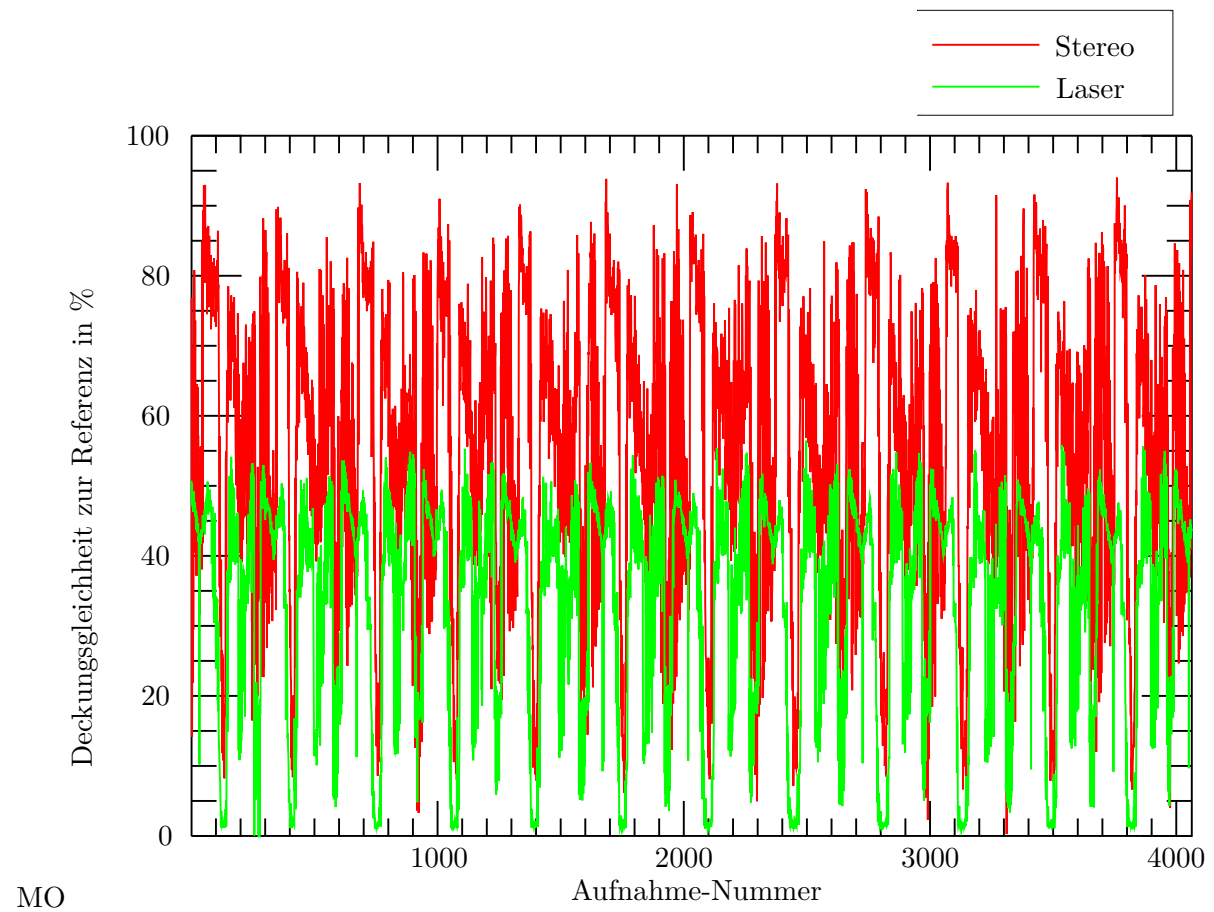
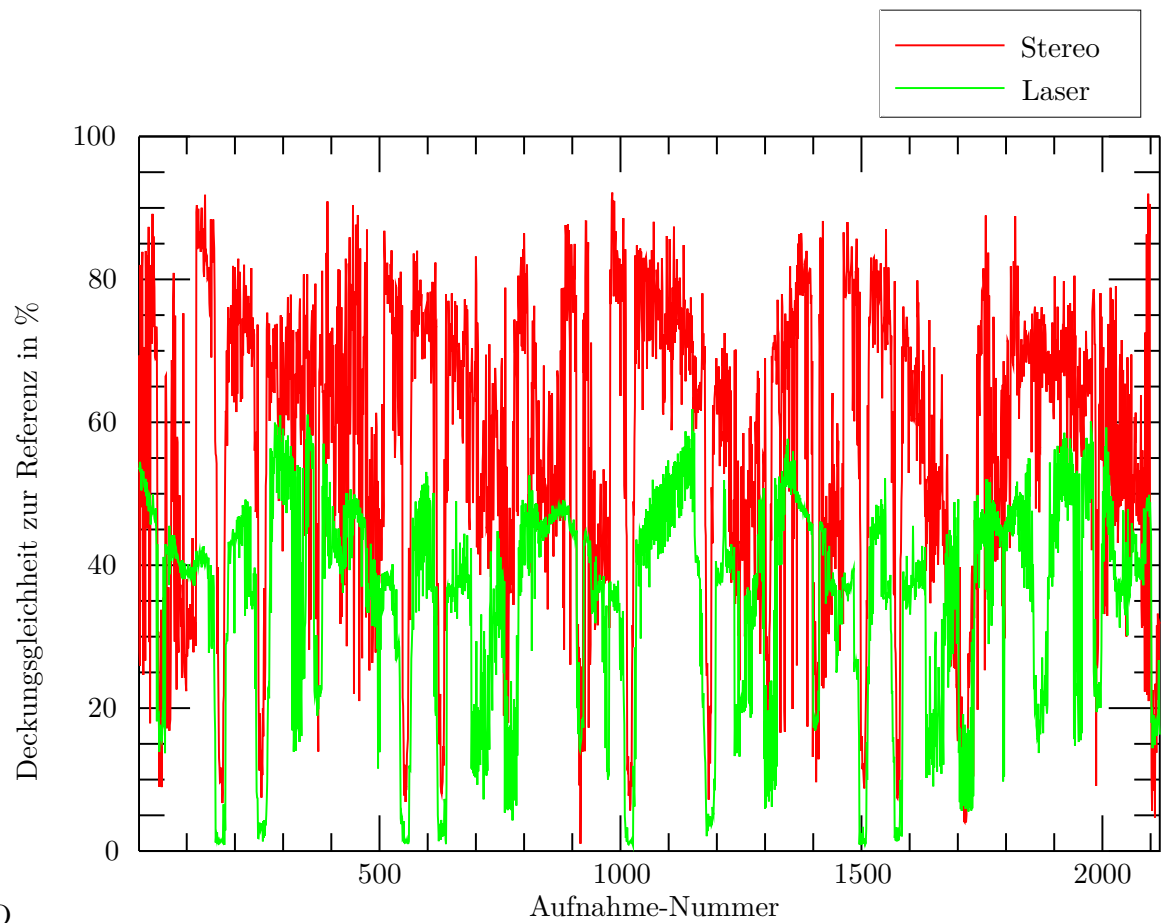


Abbildung 7.26: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den ungetrackten Messobjekten der Sensoren und den Referenzobjekten für die Route 2



MO

Abbildung 7.27: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den ungetrackten Messobjekten der Sensoren und den Referenzobjekten für die Route 3

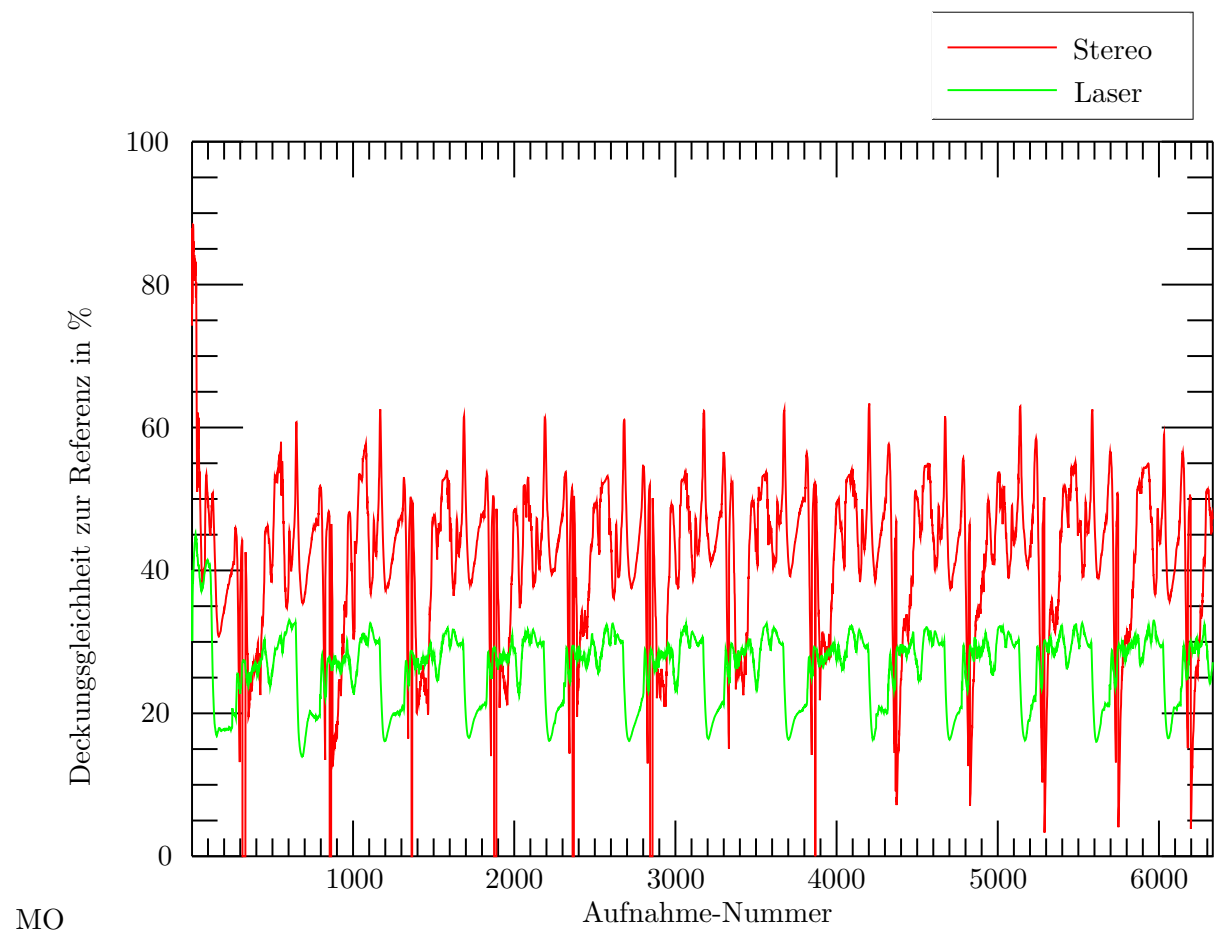


Abbildung 7.28: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der Sensoren und den Referenzobjekten für die Route 1

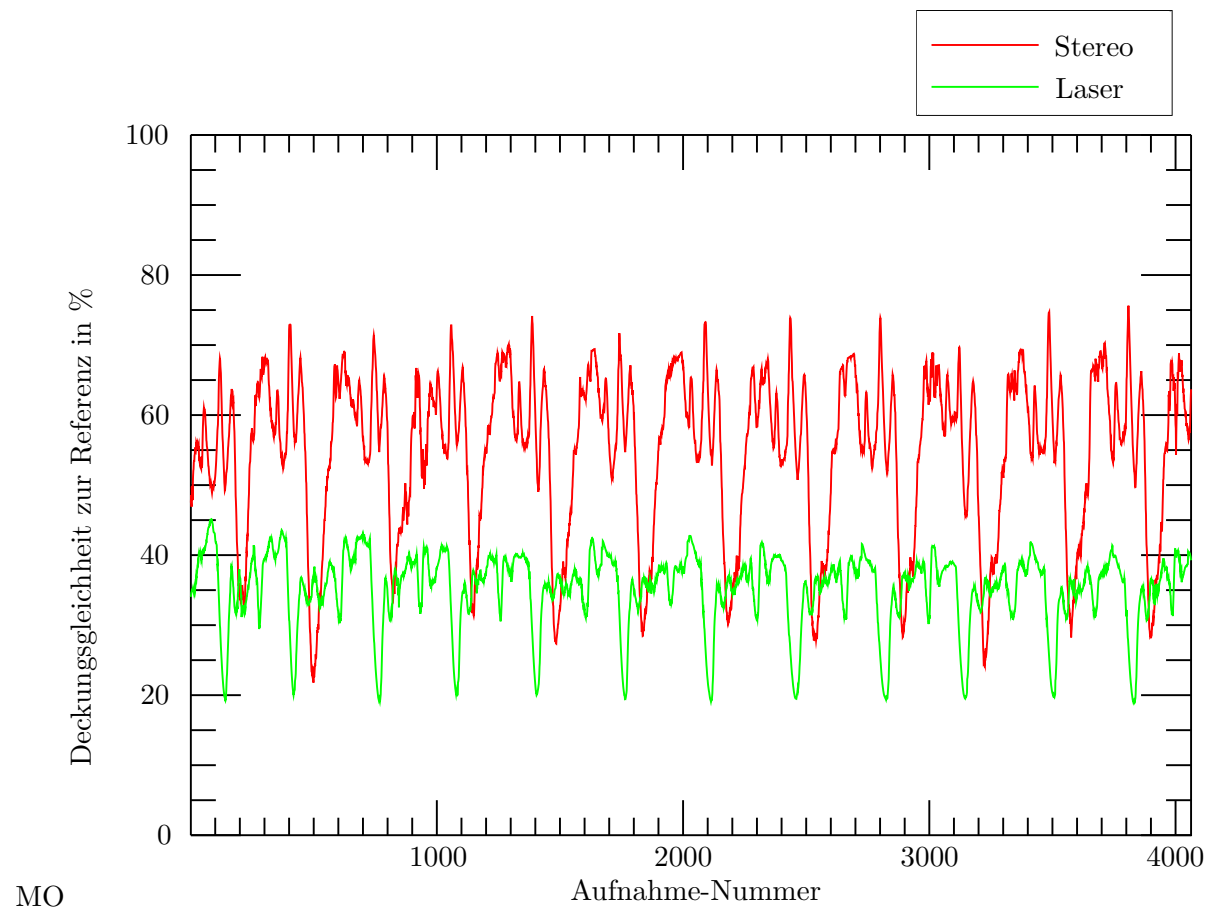


Abbildung 7.29: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der Sensoren und den Referenzobjekten für die Route 2

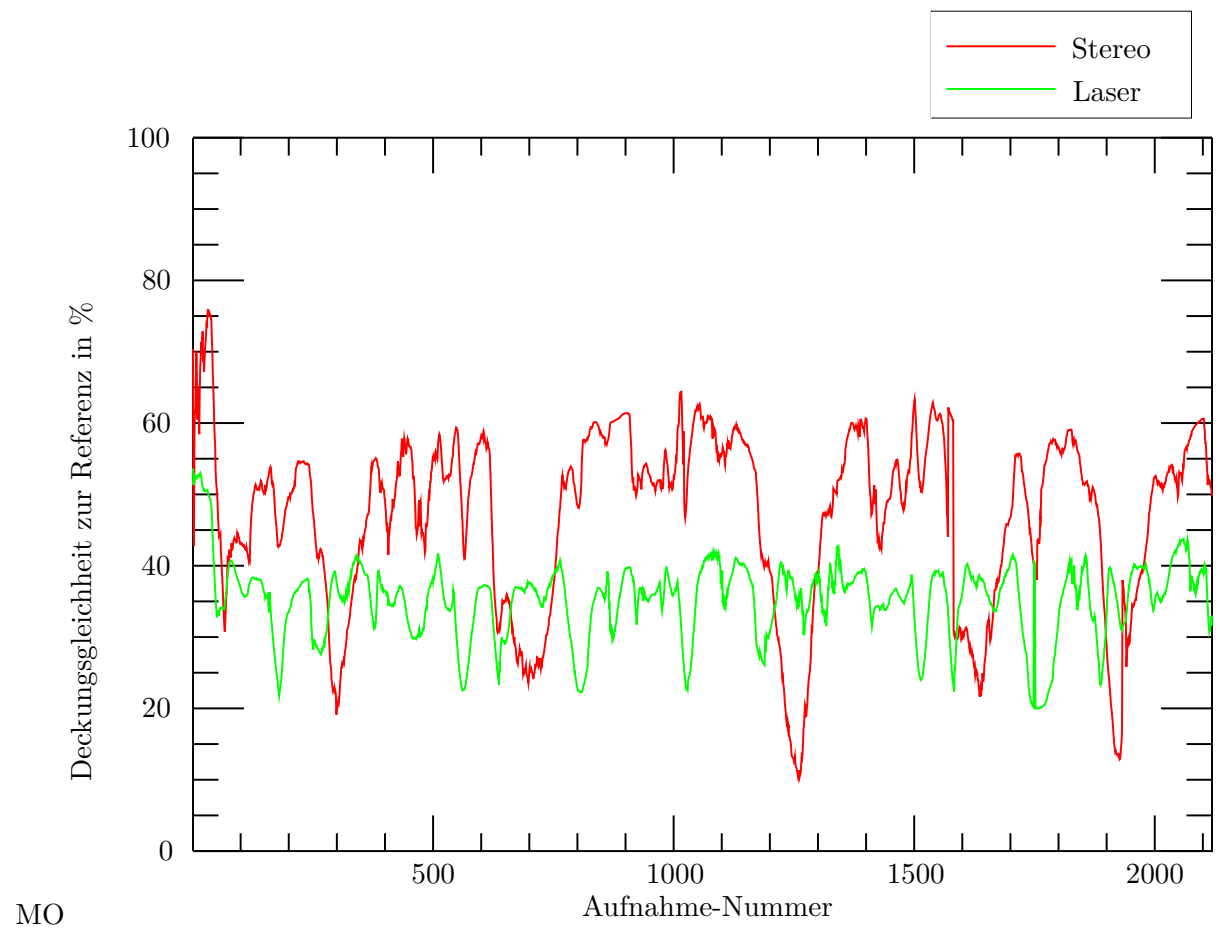


Abbildung 7.30: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der Sensoren und den Referenzobjekten für die Route 3

7.6.2 Vergleich der Fusionsebenen

Wie ebenfalls bereits in [50] gezeigt, erzielt die LLF beim direkten Vergleich mit der MLF in der jeweiligen Route eine höhere Genauigkeit in der Objektschätzung (siehe Tabelle 7.11 und Abbildungen 7.31, 7.32 und 7.33). Erwartungsgemäß ist auch im Vergleich mit den unfusionierten Sensordaten die mittlere Abweichung der Messergebnisse zu den Referenzobjekten bei der LLF am geringsten (siehe Tabelle 7.10). Die Genauigkeit der Objektbestimmung der MLF liegt mit Ausnahme von Route 1 zwischen den erzielten Genauigkeiten der unfusionierten Objektdaten der beiden Sensoren. In Route 1 sind die getrackten MLF-Objekte im Schnitt sogar genauer als die unfusionierten Daten der Sensoren.

Ähnlich zu den unfusionierten Sensordaten werden bei der Route 3 die Referenzobjekte von den getrackten Objekten schlechter abgebildet als von den nicht getrackten Objekten.

	LLF ohne Tracking	LLF mit Tracking	MLF ohne Tracking	MLF mit Tracking
Fahrt 1	16.59 m^3	16.74 m^3	24.5 m^3	21.87 m^3
Fahrt 2	11.98 m^3	11.5 m^3	20.12 m^3	18.96 m^3
Fahrt 3	11.22 m^3	12.14 m^3	18.9 m^3	19.56 m^3

Tabelle 7.11: Mittlere Abweichung der Objektbox-Volumina zwischen den Messobjekten der Sensordatenfusion und den Referenzobjekten

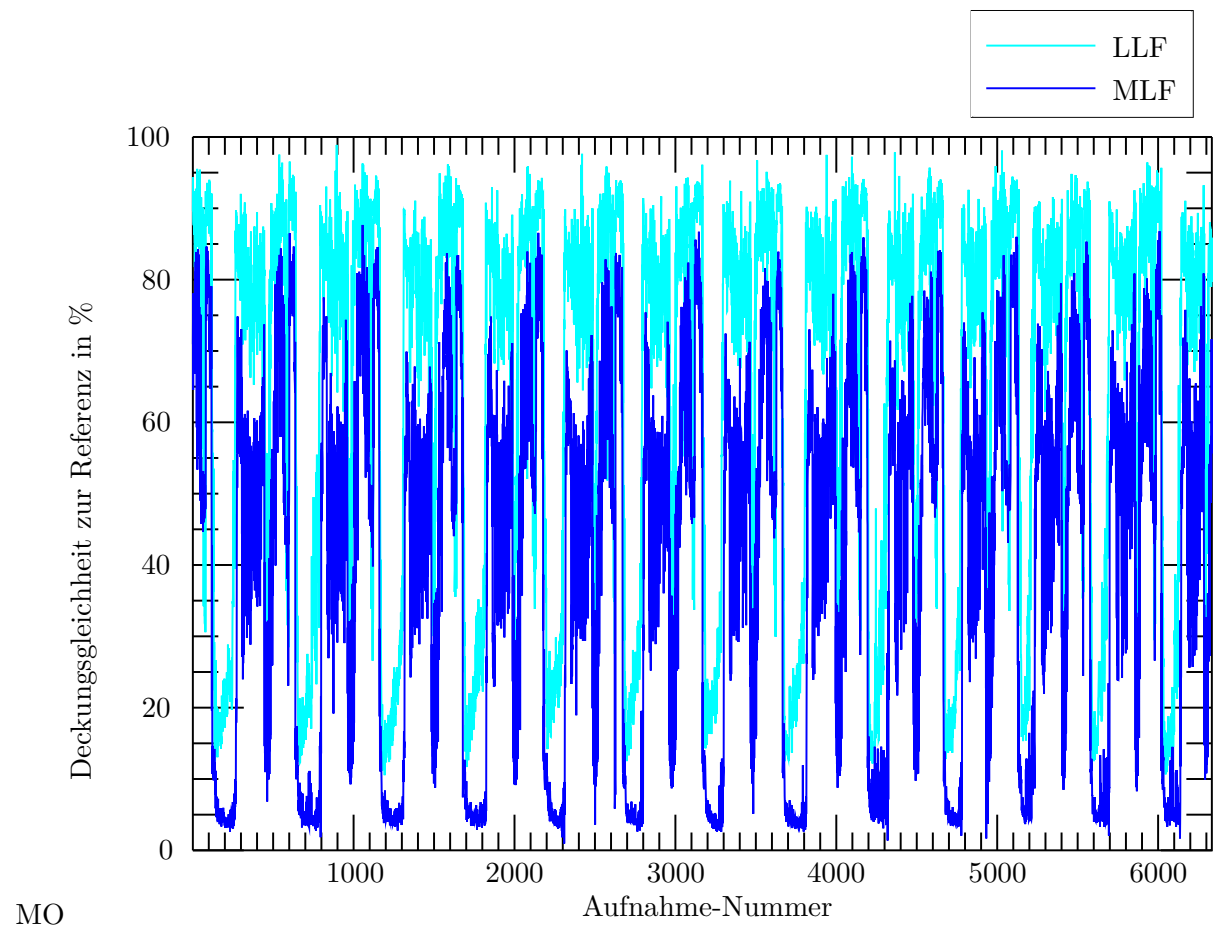


Abbildung 7.31: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den ungetrackten Messobjekten der *LLF* und *MLF* und den Referenzobjekten für die Route 1

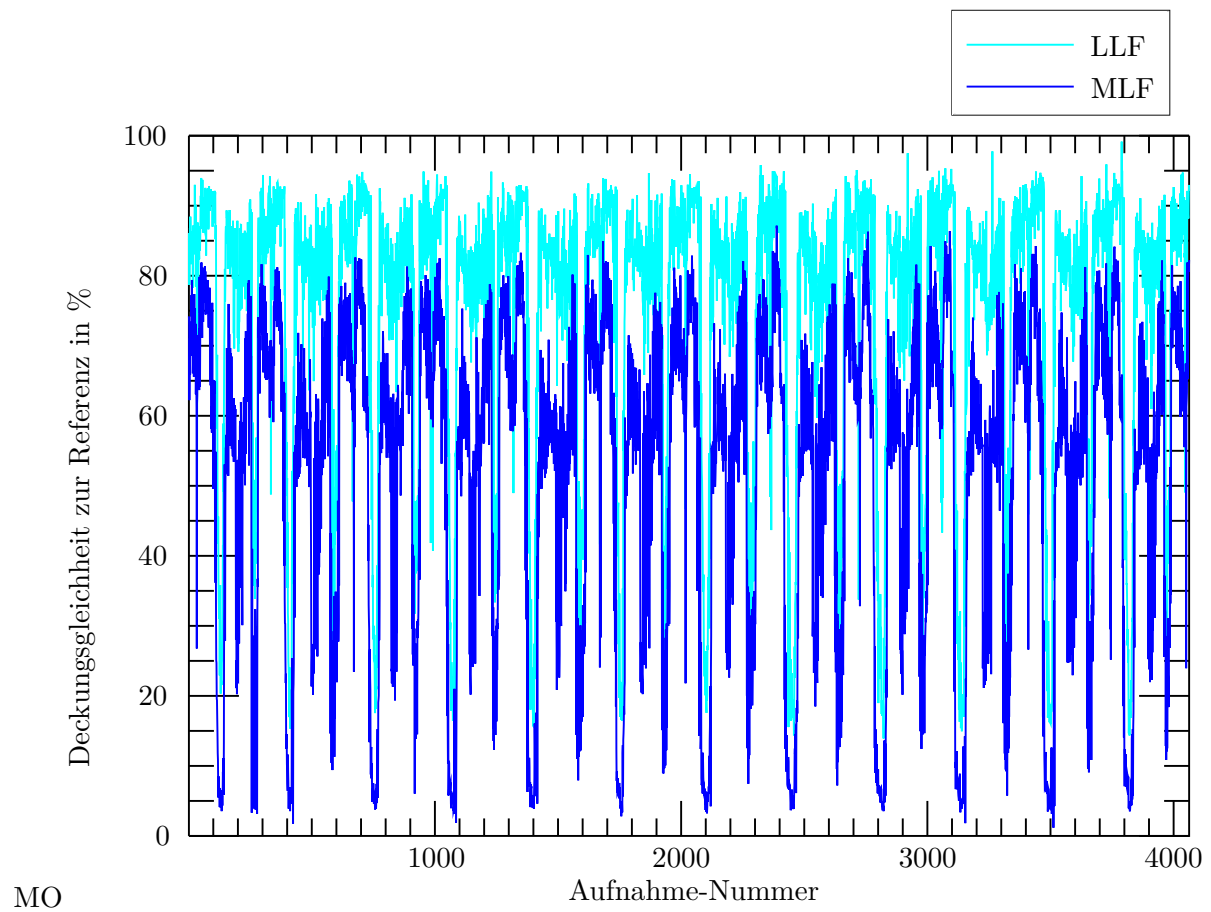


Abbildung 7.32: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den ungetrackten Messobjekten der *LLF* und *MLF* und den Referenzobjekten für die Route 2

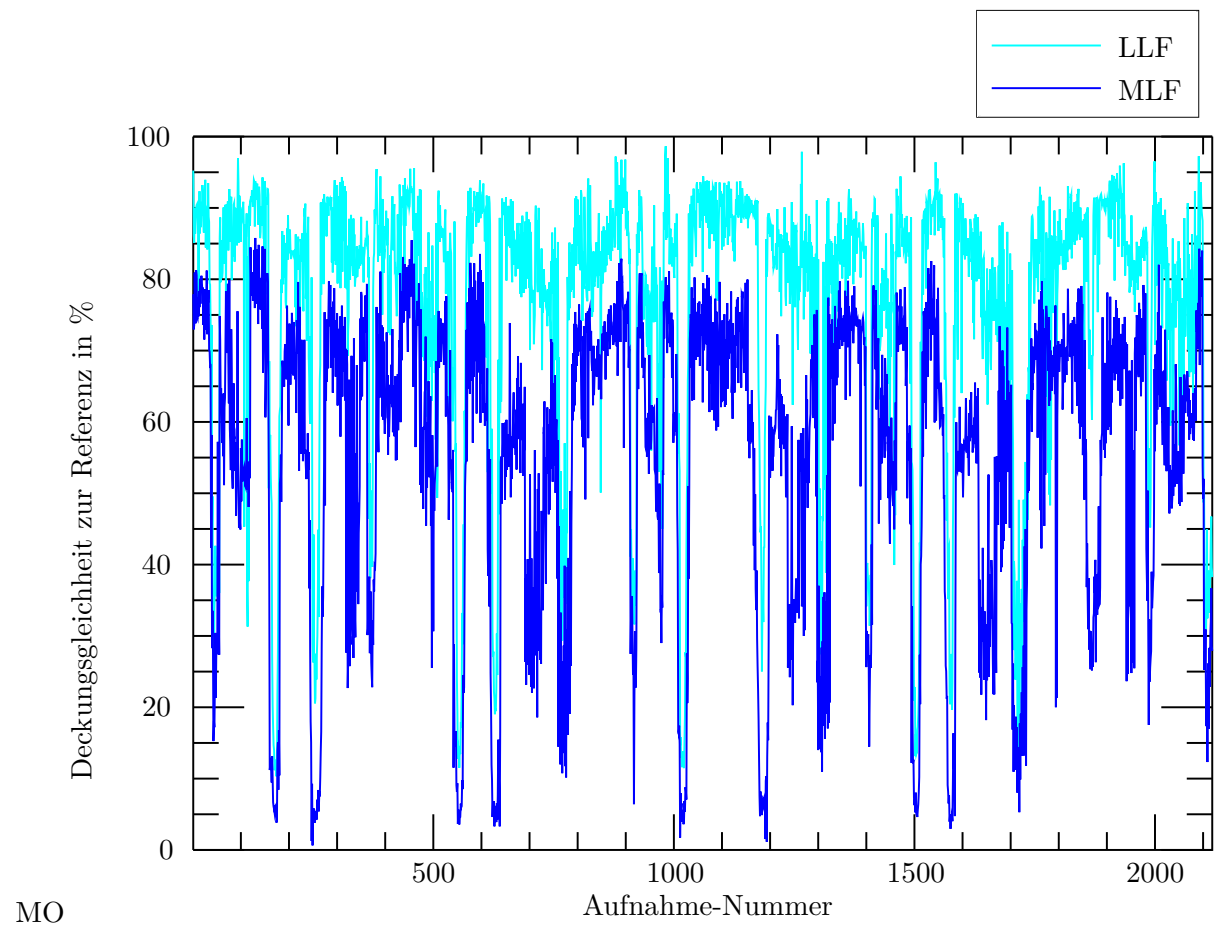


Abbildung 7.33: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den ungetrackten Messobjekten der *LLF* und *MLF* und den Referenzobjekten für die Route 3

Die getrackten Objektdaten der drei Verfahren zur HLF werden mit den getrackten Objektdaten des Referenzfahrzeuges validiert. Hierbei wird insgesamt durch die SCF im Mittel die geringste Abweichung zur Referenz erzielt (siehe Tabelle 7.12). Die IF ist nur geringfügig ungenauer oder in Route 3 sogar genauer. Die Ähnlichkeit der Ergebnisse deuten auf die enge Verwandtschaft der beiden Verfahren hin, da die IF als eine Erweiterung der SCF angesehen werden kann. Die SCF ist in allen drei gefahrenen Routen sogar genauer als die unfusionierten Objektdaten der beiden Sensoren (siehe Tabelle 7.10).

Die FCIF ist vergleichsweise deutlich ungenauer, da sie sich in der Fusion zu stark an den im Schnitt ungenaueren Laserscannerdaten orientiert. Der Grund liegt in der Verwendung des Gesamtgewichts ω (siehe Gleichungen 6.31 und 6.32), welches zusätzlich zum Genauigkeitsverhältnis zweier Fusionsobjekte die Fusion steuert (nur für $\omega_i = 0.5$ entspricht FCIF der SCF). Anders als bei der IF und SCF wird ω aus allen Zustandsgrößen gleichsam ermittelt und wirkt sich anschließend auf alle Zustandsgrößen in der Fusion gleich stark aus, was zu einer ungerechten Gewichtung einzelner Größen führen kann. Bei der vorliegenden Messung wird somit die Tendenz in der Fusion zu den Laserscanner-Objekten verstärkt.

	HLF-SCF	HLF-IF	HLF-FCIF
Fahrt 1	21.78 m^3	22.31 m^3	27.68 m^3
Fahrt 2	17.49 m^3	17.69 m^3	24.41 m^3
Fahrt 3	17.76 m^3	17.69 m^3	23.22 m^3

Tabelle 7.12: Mittlere Abweichung der Objektbox-Volumina zwischen den Messobjekten der drei unterschiedlichen Verfahren zur High-Level-Fusion und den Referenzobjekten

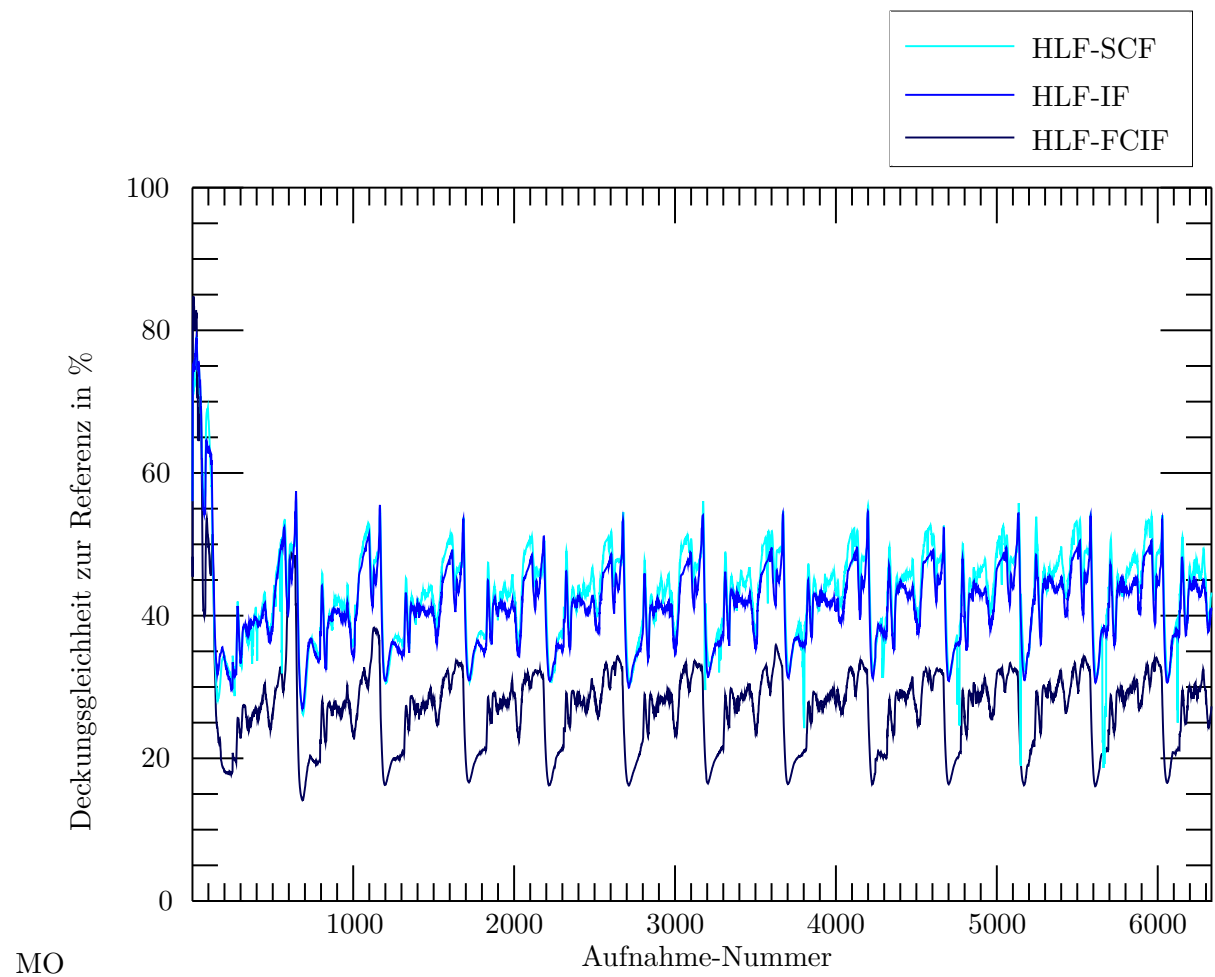


Abbildung 7.34: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der drei Verfahren zur HLF (*SCF*, *IF*, *FCIF*) und den Referenzobjekten für die Route 1

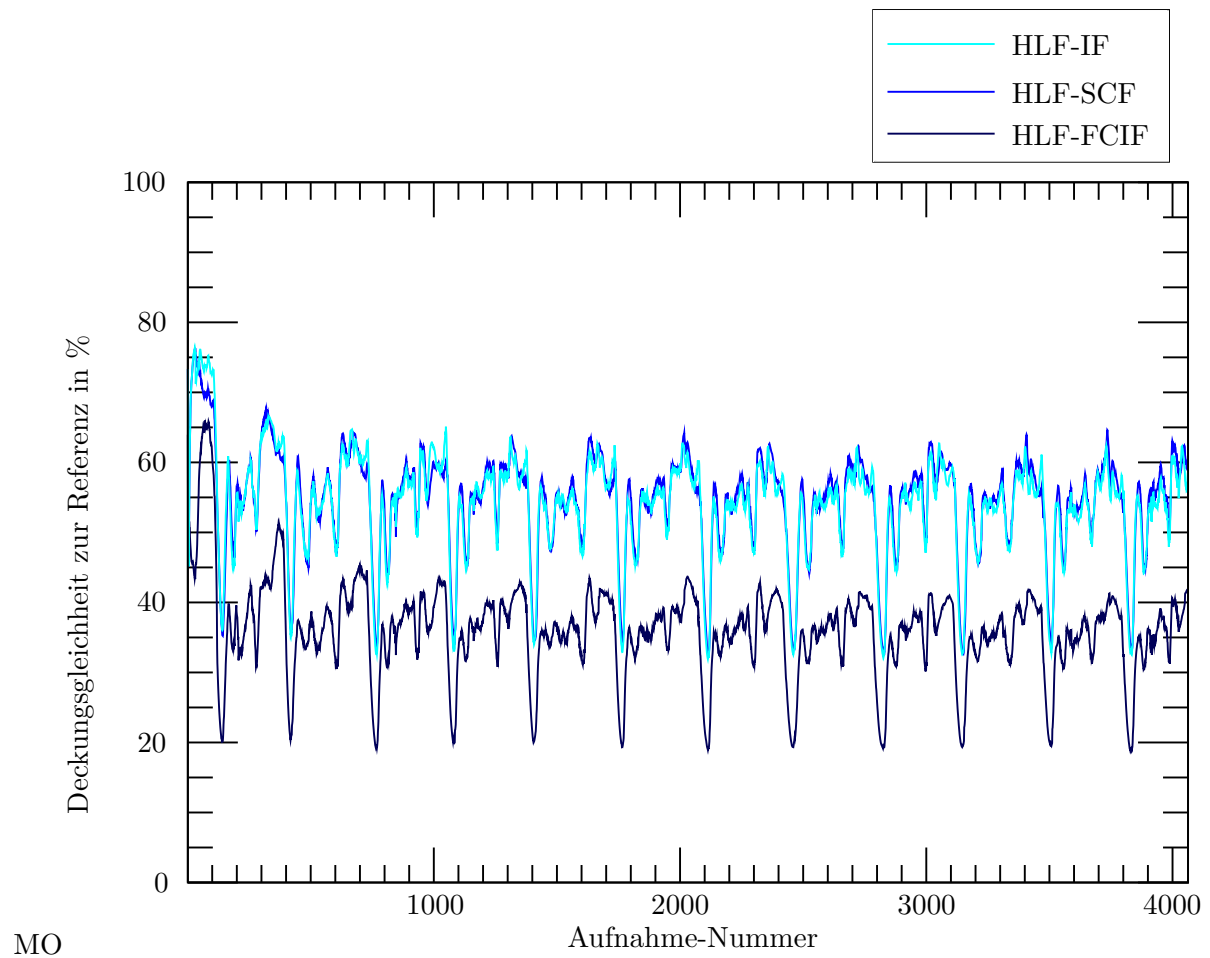


Abbildung 7.35: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der drei Verfahren zur HLF (*SCF*, *IF*, *FCIF*) und den Referenzobjekten für die Route 2

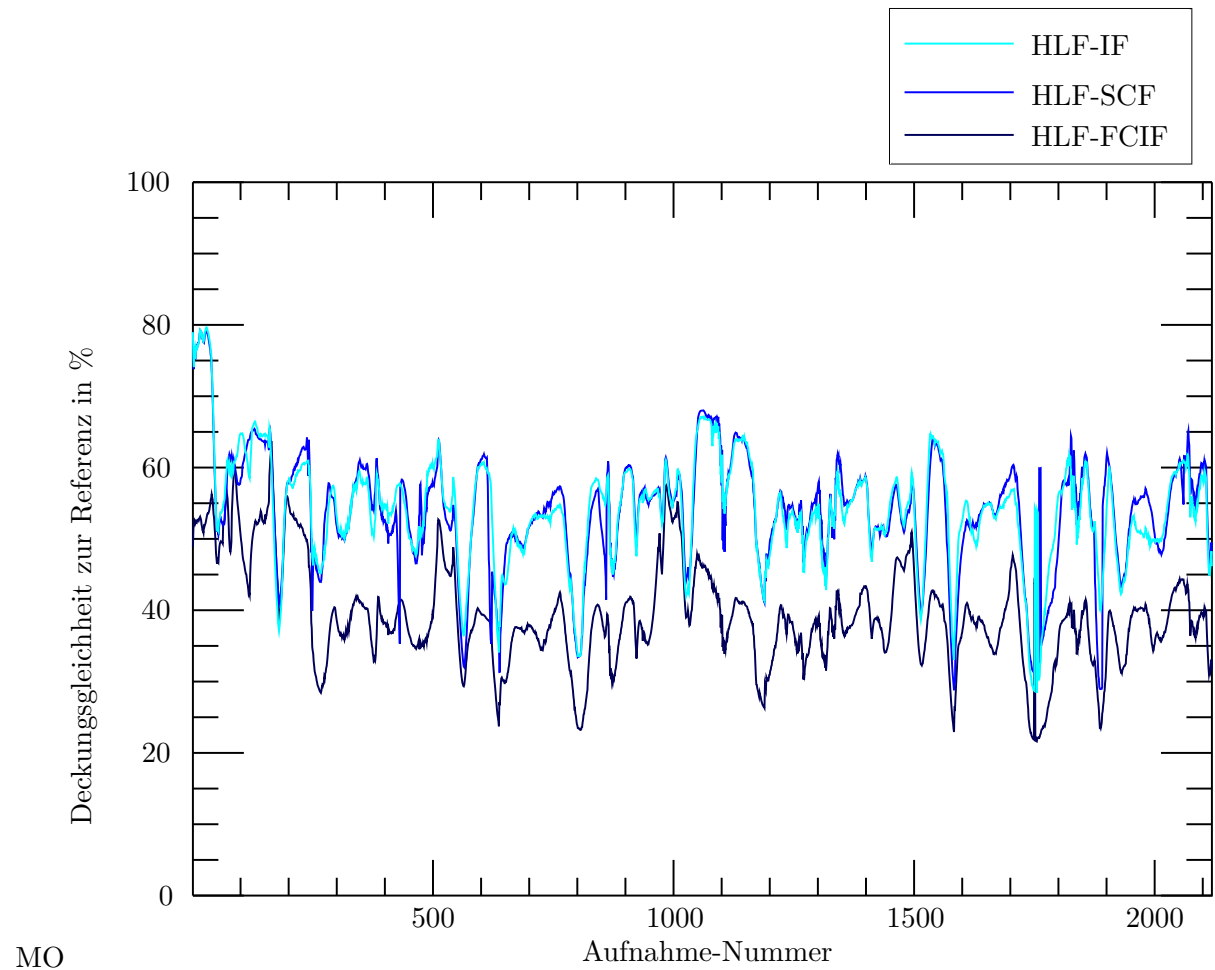


Abbildung 7.36: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der drei Verfahren zur HLF (*SCF*, *IF*, *FCIF*) und den Referenzobjekten für die Route 3

Werden die getrackten Objektdaten der drei Fusionsebenen gegenüber gestellt, so erreicht die LLF erwartungsgemäß in allen drei gefahrenen Routen die höchste Überdeckung mit den Referenzwerten (siehe Abbildungen 7.37, 7.38 und 7.39). Etwas unerwartet erzielt die HLF, zumindest mit der SCF und IF, eine höhere Genauigkeit als die MLF. Eine direkte Vergleichbarkeit von MLF und HLF ist dadurch verstärkt gewährt, da die Zustandsfusion von MLF und SCF verfahrensgleich ist. Besonders in der Auswertung zur Route 1 und 3 (siehe Abbildungen 7.37 und 7.39) ist zu erkennen, dass anfangs die beiden Verfahren noch relativ ähnliche Genauigkeiten erzielen und erst danach ein erkennbarer Offset entsteht. Zudem weisen die Kurvenverläufe eine hohe Ähnlichkeit auf. Eine mögliche Ursache könnte die Schätzung der Objektdimensionen sein, die ohne ein entsprechendes Modell im EKF gefiltert werden und durch das mehrmalige Tracking in der HLF begünstigt werden.

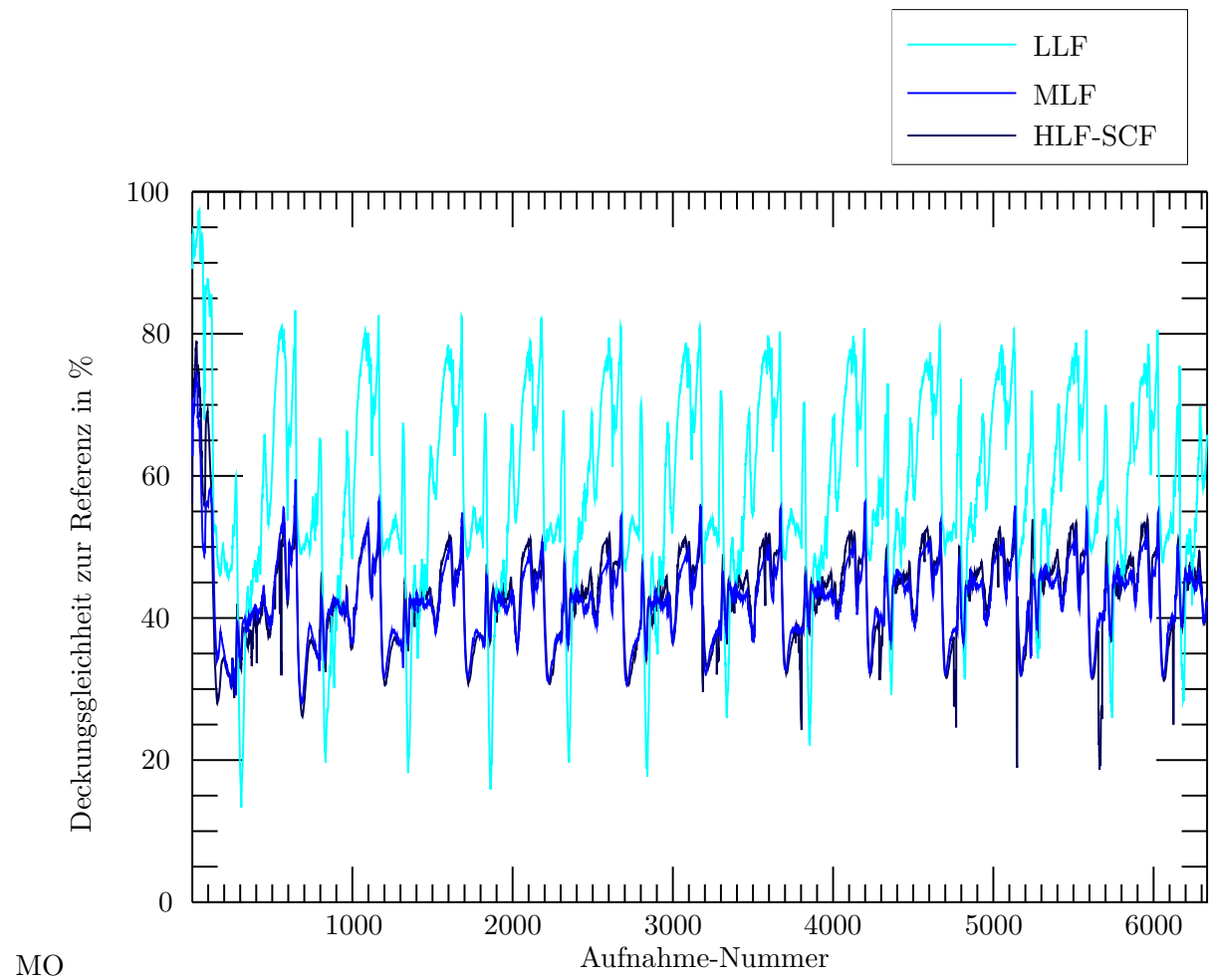


Abbildung 7.37: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der *LLF*, *MLF* und *HLF-SCF* und den Referenzobjekten für die Route 1

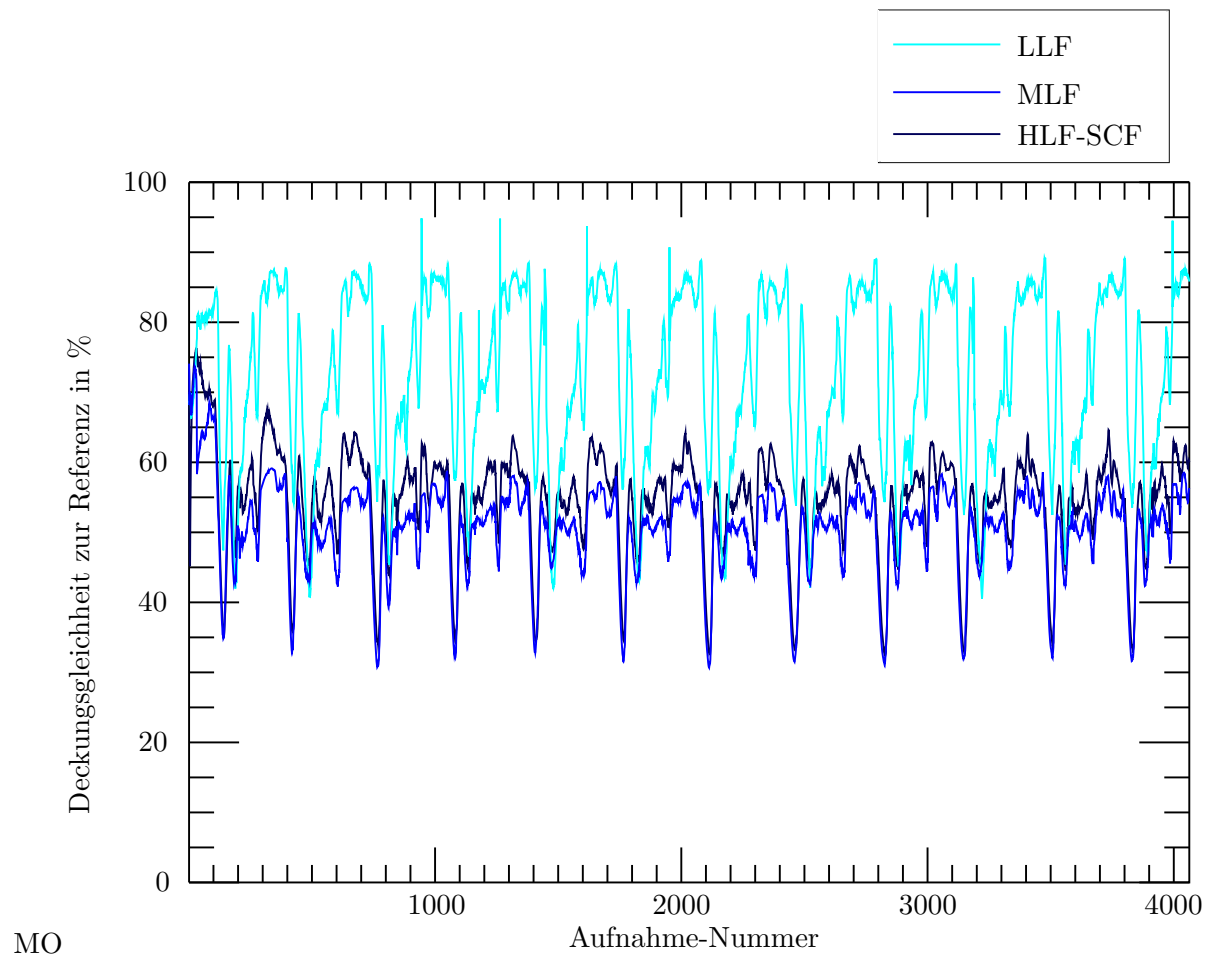


Abbildung 7.38: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der *LLF*, *MLF* und *HLF-SCF* und den Referenzobjekten für die Route 2

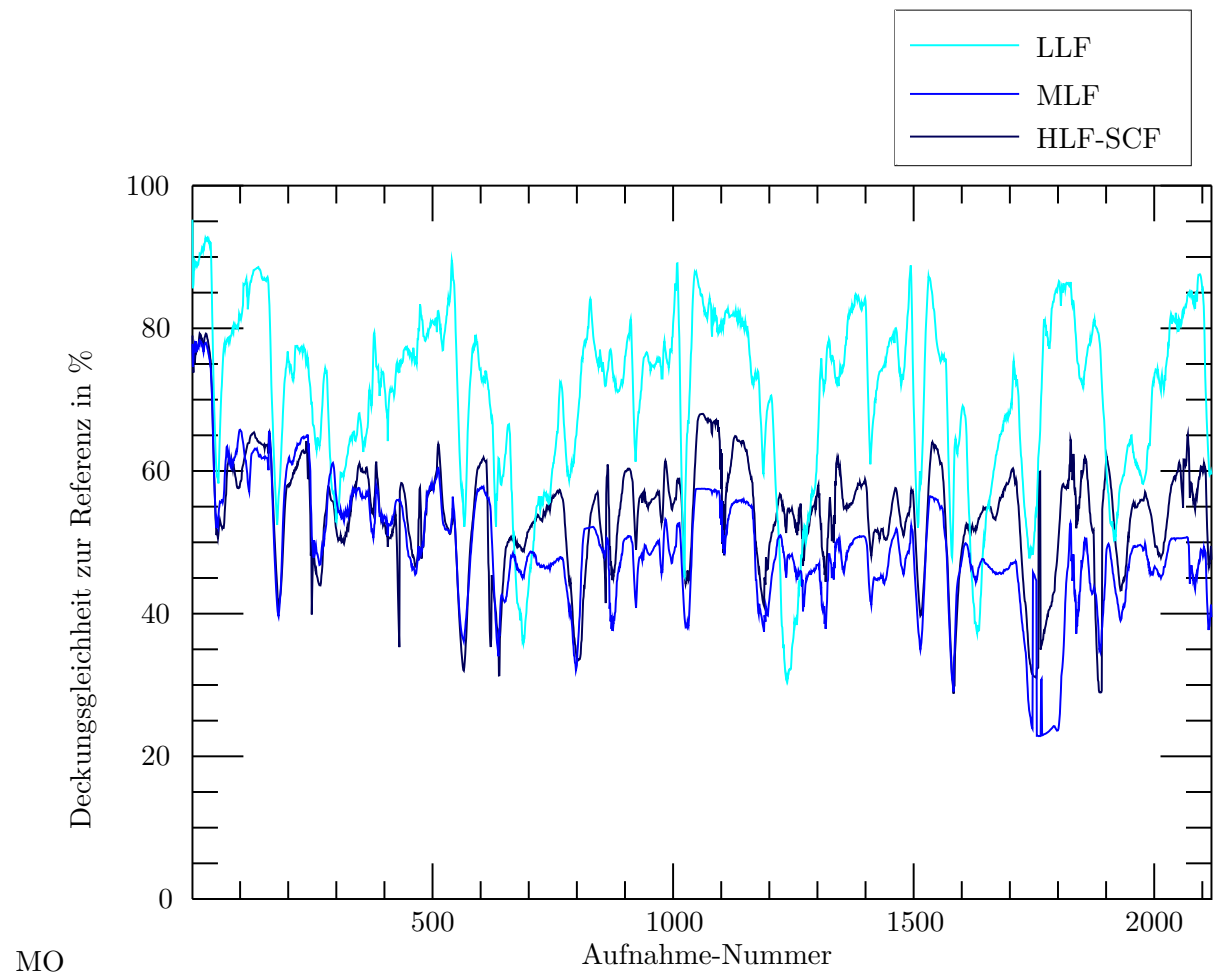


Abbildung 7.39: Darstellung der prozentualen Überdeckung der Objektbox-Volumina zwischen den getrackten Messobjekten der *LLF*, *MLF* und *HLF-SCF* und den Referenzobjekten für die Route 3

7.7 Zusammenfassung

Im vorangehenden Kapitel wurden Auswertungsverfahren und deren Ergebnisse zur Überprüfung der vorgestellten Methoden zusammengefasst. Die genauere Objektbeschreibung des Trackings mit Referenzpunktwechsel im Vergleich zum Tracking des Mittelpunktes wurde in einer Simulation für den Fall einer eingeschränkten Beobachtbarkeit aller Referenzpunkte am Objekt nachgewiesen. Aufgrund mangelnder realer Messdaten wurden die Verfahren zur Bestimmung und Beseitigung der Mess- und Verarbeitungsfehler ebenfalls simulativ überprüft. Bei der Konfliktauflösung zur Fehl- und Falschdetektion bzw. der Fragmentierung und Zusammenfassung eines Objektes wurde die Abhängigkeit der Verfahren von der lokalen Messgenauigkeit und dem Segmentierungsverhalten der einzelnen Sensoren dargestellt. Durch den direkten Vergleich der lokalen Detektionspotentiale der einzelnen Sensoren wird der Konflikt dynamisch aufgelöst. Bei mehr als zwei beobachtenden Sensoren geht zusätzlich das mengenmäßige Verhältnis der komplementären Hypothesen als Gewichtung in die Konfliktauflösung mit ein. Bei der Untersuchung der MOT-Erkennung durch zwei Simulationsszenarien hat sich gezeigt, dass schon die widersprüchlichen Daten zweier Sensoren mit vergleichbarer Messgenauigkeit ausreichen können, um den Fehler sicher zu erkennen. Die Genauigkeit der Objektbestimmung wurde in einem realen Testszenario mit hochgenau vermessenen Bewegungen eines Referenzfahrzeugs ermittelt. Hierbei erzielte die LLF die höchste Genauigkeit im Vergleich zu den beiden einzelnen Sensoren sowie zur MLF und HLF. Bei den Verfahren zur Zustandsfusion der Track-Daten ergaben sich durch die SCF und IF nahezu identische Genauigkeiten. Die FCIF schnitt bei der Objektschätzung signifikant schlechter ab.

8 Zusammenfassung und Ausblick

Das maßgebliche Ziel der Arbeit war die Entwicklung und Untersuchung einer mehrschichtig gekoppelten Sensorfusionsarchitektur zur Reduzierung der fehlerhaften Ausgaben und zur Steigerung der Genauigkeit und Zuverlässigkeit in der sensorgestützten Fahrzeugumfelderfassung. Hierfür wurde eine bis dato weitestgehend neuartige Sensorkombination eines Laserscanners und einer Stereokamera als Referenzsystem aufgebaut und Verfahren zur Kalibrierung und Datenprozessierung entwickelt (Kapitel 2-5). Die konkurrierenden Sensorinformationen wurden auf den drei Verarbeitungsebenen durch neu entwickelte, erweiterte und bewährte Verfahren zusammengefasst und die Resultate entsprechend validiert (Kapitel 6-7).

Kapitel 2 gibt einen Überblick zum Stand der Technik in der sensorgestützten Fahrzeugumfelderfassung. Die existierenden Ansätze zur Datenfusion von homogenen und heterogenen Sensorsystemen, insbesondere in Hinblick auf den speziellen Anwendungsfall der Fahrzeugumfelderfassung, werden vorgestellt und miteinander verglichen. Im Zuge dessen wird die neu entwickelte Multisensor-Multilevel-Fusion eingeführt und den existierenden Verfahren gegenübergestellt.

In Kapitel 3 wurden die Verfahren zur Datenvorverarbeitung und Sensorkalibrierung vorgestellt. Hervorzuheben ist der innovative Ansatz zur gleichzeitigen Bestimmung der äußeren Orientierung zum SKS und des systematischen Fehleranteils des Laserscanners durch die Verwendung von Ebenen als Kalibrierkörper. Zusätzlich wurden Ansätze zur Bestimmung des zufälligen Fehleranteils für beide Sensoren neu entwickelt oder bestehende Verfahren erweitert. So wurde beispielsweise beim Stereosystem neben den bekannten dreistufigen Orientierungsverfahren der Einfluss des Stereo-Matching-Verfahrens auf den zufälligen Fehler in der Punktbestimmung mit einbezogen.

Die Extraktion der Segment- und Objektinformationen aus den räumlichen Punktdaten wurde in Kapitel 4 behandelt. Das entwickelte Segmentierungsverfahren wird durch die punktweise Bestimmung von Segmentierungsverhalten und Messgenauigkeit in Abhängigkeit zum Sensortyp gesteuert. Durch die Entkopplung der Segmentierung vom sensorspezifischen Segmentierungsverhalten und dem Messfehler gibt es hierbei keine Eingrenzung auf bestimmte Sensorverfahren. Auf Grundlage der Segmentierungs- und Fehlermodellierung wurde zudem ein analytischer Ansatz zur Schätzung der Segmentierungswahrscheinlichkeit als wichtige Basis der anschließenden Datenverarbeitung und Fusion geliefert. Zur umfassenden und robusten Bestimmung der Objektboxausrichtung in der Horizontalen wurde das bestehende Zwei-Linien-Verfahren um ein Verfahren zur Ellipsenschätzung erweitert.

Die räumliche Verfolgung der Objekte durch das Tracking wurde in Kapitel 5 formuliert. Das Prädiktions-Korrektur-Konzept wurde mittels eines EKF umgesetzt, der neben dem aktuellen Referenzpunkt die Objektausmaße und Objektausrichtung filtert. Zur rechnerischen Optimierung der Objektassoziation wurde das Verfahren der Assoziationsmatrix unter Einbeziehung von Entfernungs- und Ähnlichkeitsmerkmalen mit einem GNN-Verfahren neuartig kombiniert. Des Weiteren wurde zur Vermeidung von Verzerrungen im Tracking-Prozess, durch einen zu un-

genau bestimmten Referenzpunkt, ein Beobachtungsmodell entwickelt. Hierdurch ist es möglich bei einem Wechsel des optimal beobachtbaren Referenzpunktes den getrackten Objektzustand entsprechend anzupassen. Der Vorteil gegenüber dem Tracking am Objektmittelpunkt wurde in der Simulation verdeutlicht.

Die Verfahren zur Datenfusion auf den drei Prozessierungsebenen wurden in Kapitel 6 vorgestellt. Die einzelnen Verfahren sind dabei unabhängig vom Typ und Anzahl der eingehenden Sensoren. Die Low-Level-Fusion entspricht dem entwickelten Segmentierungsprozess für mehrere eingehende Sensoren. Bei der Mid-Level-Fusion wurde neben der Fusion der Objektparameter Verfahren zur wahrscheinlichkeitsbasierten Auflösung widersprüchlicher Sensoraussagen bezüglich Fehl- und Falschdetektionen (FPO/FNO-Test) und Objektzusammenfassungen und -fragmentierungen (ZO/FO-Test) entwickelt. Analog wurden auf der Track-Ebene bei der High-Level-Fusion Sensorwidersprüche hinsichtlich Fehl- und Falschdetektionen (FPT/FNT-Test) und Objektwechsel innerhalb von Tracks (MOT-Test) aufgelöst und die Funktionsweise in der Simulation überprüft. Als Grundlage der Konflikterkennung dienen die Ergebnisse einer neu eingeführten Objekt- und Track-Clusterung. Beim direkten Vergleich von drei untersuchten Verfahren zur Track-Fusion durch ein reales Testszenario erreichten die SCF und IF die höchsten Genauigkeiten in der Objektschätzung, gefolgt von der FCIF. Die SCF und IF erzielten sogar eine durchschnittlich höhere Genauigkeit als die MLF. Die beste Gesamtgenauigkeit in der Objektschätzung der drei Fusionsebenen und der beiden einzelnen Sensoren erreichte die LLF. Abschließend wurde das vorgestellte MMFF über eine Kopplung der Fusionsverfahren erweitert, um die hohe Genauigkeit der LLF in der Objektbestimmung mit den Verfahren zur Aufdeckung und Bereinigung der Detektions- und Verarbeitungsfehler der oberen beiden Fusionsebenen zu kombinieren.

Die Hauptinnovationen der Arbeit stellen die vom Messverfahren und der Messgenauigkeit gesteuerte Segmentierung und Low-Level-Fusion sowie die neuen Ansätze zur modellgesteuerten Erkennung und Entfernung der Mess- und Verarbeitungsfehler durch redundante und konkurrierende Sensorinformationen auf den unterschiedlichen Fusionsebenen dar. Eine Zusammenführung dieser beiden Neuerungen wird durch eine innovative Kopplung der Fusionsebenen im vorgestellten hierarchischen MMFF erzielt.

Die in der Einleitung aufgestellten Forschungsthemen können rückblickend folgendermaßen beurteilt werden:

- **Optimale Zustandsbestimmung der Objekte durch Low-Level-Fusion**
 - Die Auswertung der Experimente hat gezeigt, dass im Vergleich zur MLF, HLF und den Einzelsensoren durch die LLF die höchste Genauigkeit bei der Bestimmung der Objektparameter erzielt wird.
- **Erkennung und Entfernung von Mess- und Prozessierungsfehlern durch Fusion auf höheren Verarbeitungsebenen**
 - Es wurden Verfahren zur Erkennung und Bereinigung der kategorisierten Prozessierungsfehler auf der mittleren und oberen Fusionsebene eingeführt und deren Funktionsweise in der Simulation überprüft und dargestellt.

- **Reduzierung von Prozessierungsfehlern bei gleichzeitig optimaler Objektbestimmung durch eine kombinierte Fusion auf verschiedenen Prozessierungsebenen**

- Das vorgestellte MMFF ist neben der Datenfusion auf den drei unterschiedlichen Verarbeitungsebenen in der Lage über eine Kopplung der Prozessierungsebenen die Fehlerauflösung der höheren Ebenen mit der Objektbestimmung der unteren Ebene zu verknüpfen.

Neu gewonnenes Wissen stellt wiederum den Ausgangspunkt für eine weiterführende Forschung dar. Für den anschließenden Ausbau des MMFF und zur weiteren Validierung der Verfahren sind die hier teilweise synthetisch erzeugten Daten durch zusätzliche reelle Messdaten zu ergänzen oder zu ersetzen. Für diese Arbeit konnte ein umfassendes Messszenario, bedingt durch den experimentellen Aufwand, nicht mehr realisiert werden.

Die vorgestellte Kompensation der beim Stereosystem aufgetretenen Modell-Realitätsabweichung in der Bestimmung der SW durch eine manuelle Kalibrierung könnte durch eine entsprechende Modellerweiterung substituiert werden.

Ein weiterer interessanter Aspekt liegt in der Einbindung von zusätzlichem Infrastrukturwissen, beispielsweise Karteninformationen, zur Stützung der Auflösung von Fehl- und Falschdetektionen.

In Hinblick auf den zu erwartenden Einsatz von drahtlosen Kommunikationstechnologien zwischen Fahrzeugen und der Infrastruktur erscheint eine Erweiterung des MMFF-Frameworks zur Aufnahme von externen Sensorinformationen von der Infrastruktur (*Car-2-Infrastructure*) bzw. von benachbarten Fahrzeuge (*Car-To-Car*) in die Datenfusion als sinnvoll.

Literaturverzeichnis

- [1] ALLESANDRETTI, G. ; BROGGI, A. ; CERRI, P. : Vehicle and Guard Rail Detection Using Radar and Vision Data Fusion. In: *IEEE Transactions on Intelligent Transportation Systems* Bd. 8, 2007, S. 95–105
- [2] AMDITIS, A. ; IURGEL, U. : From Fusion to Perception : The Perception Subproject of INTERACTIVE. In: *8th ITS European Congress Lyon*, 2011
- [3] BAIG, Q. ; VU, T.-D. ; AYCARD, O. : Online localization and mapping with moving objects detection in dynamic outdoor environments. In: *Proc. IEEE 5th Int. Conf. Intelligent Computer Communication and Processing ICCP 2009*, 2009, S. 401–408
- [4] BAR-SHALOM, Y. : On the track-to-track correlation problem. 26 (1981), Nr. 2, S. 571–572. <http://dx.doi.org/10.1109/TAC.1981.1102635>. – DOI 10.1109/TAC.1981.1102635
- [5] BAR-SHALOM, Y. ; FORTMANN, T. E.: *tracking and data association*. Bd. 179. Academic, 1988
- [6] BAR-SHALOM, Y. ; TSE, E. : Tracking in a Cluttered Environment with Probabilistic Data Association. In: *Automatica*, 1975, S. 451–460
- [7] BAR-SHALOM, Y. ; CAMPO, L. : The Effect of the Common Process Noise on the Two-Sensor Fused-Track Covariance. (1986), Nr. 6, S. 803–805. <http://dx.doi.org/10.1109/TAES.1986.310815>. – DOI 10.1109/TAES.1986.310815
- [8] BEDWORTH, M. ; O'BRIEN, J. : The Omnibus model: a new model of data fusion? In: *IEEE Aerospace and Electronic Systems Magazine* 15 (2000), Nr. 4, S. 30–36. <http://dx.doi.org/10.1109/62.839632>. – DOI 10.1109/62.839632
- [9] BERNARDO, J. ; SMITH, A. : *Bayesian Theory*. Wiley, 1994
- [10] BLACKMAN, S. ; POPOLI, R. : *Design and Analysis of Modern Tracking Systems*. Artech House, 1999
- [11] BLOM, H. A. P.: An efficient filter for abruptly changing systems. In: *Proc. 23rd IEEE Conf. Decision and Control* Bd. 23, 1984, S. 656–658
- [12] BOEHRINGER, F. : *Gleisselektive Ortung von Schienenfahrzeugen mit bordautonomer Sensorik*, Universität Karlsruhe, Diss., 2008
- [13] BOUZOURAA, M. E. ; HOFMANN, U. : Fusion of Occupancy Grid Mapping and Model Based Object Tracking for Driver Assistance Systems using Laser and Radar Sensors. In: *IEEE Intelligent Vehicles Symposium*, 2010, S. 294–300

- [14] BOWMAN, C. L.: The dual node network (DNN) data fusion & resource management (DF & RM) architecture. In: *AIAA Intelligent Systems Conference*. Chicago, 20-22 Sep. 2004
- [15] BROGGI, A. ; CATTANI, S. ; PORTA, P. P. ; ZANI, P. : A Laserscanner-Vision Fusion System Implemented on the TerraMax Autonomous Vehicle. In: *Proc. IEEE/RSJ Int Intelligent Robots and Systems Conf*, 2006, S. 111–116
- [16] BROWN, D. : Close-range camera calibration. In: *Photogrammetric Engineering*, 1971, S. 855–866
- [17] BROWN, G. R. ; HWANG, Y. C. P.: *Introduction to random signals and applied Kalman filtering: With MATLAB exercises and solutions*. 3. New York : Wiley, 1997
- [18] CATALA PRAT, A. : *Sensordatenfusion und Bildverarbeitung zur Objekt und Gefahrenerkennung*, Technischen Universität zu Carolo-Wilhelmina Braunschweig, Diss., 2010
- [19] CHANG, K. C. ; ZHI, T. ; SAHA, R. K.: Performance evaluation of track fusion with information matrix filter. 38 (2002), Nr. 2, S. 455–466. <http://dx.doi.org/10.1109/TAES.2002.1008979>. – DOI 10.1109/TAES.2002.1008979
- [20] CHAUDHURI, D. ; SAMAL, A. : A simple method for fitting of bounding rectangle to closed regions. In: *Pattern Recognition* Bd. 40, 2007, S. 1981–1989
- [21] CHIU, C. C. ; CHUNG, M. L. ; CHEN, W. C.: Real-Time Front Vehicle Detection Algorithm for Asynchronous Binocular System. In: *Journal of Information Science and Engineering* (2010), Nr. 26, S. 735–752
- [22] CHONG, C. Y.: Hierarchical Estimation. In: *Proc. MIT/ONR Workshop on C3*, 1979
- [23] CHONG, C. Y. ; MORI, S. ; BARKER, W. H. ; CHANG, K. C.: Architectures and algorithms for track association and fusion. In: *IEEE AES Syst. Mag.*, 2000, S. 5 –13
- [24] CHONG, C.-Y. ; CHANG, K.-C. ; MORI, S. : Distributed Tracking in Distributed Sensor Networks. In: *Proc. American Control Conf*, 1986, S. 1863–1868
- [25] DANG, T. : *Kontinuierliche Selbstkalibrierung von Stereokameras*, Universität Karlsruhe, Diss., 2007
- [26] DASARATHY, B. V.: Sensor fusion potential exploitation-innovative architectures and illustrative applications. 85 (1997), Nr. 1, S. 24–38. <http://dx.doi.org/10.1109/5.554206>. – DOI 10.1109/5.554206
- [27] DEUFLHARD, P. ; HOHMANN, A. : *Numerische Mathematik 1: eine algorithmisch orientierte Einführung*. 3. Auflage. Walter de Gruyter, 2002
- [28] DINGUS, T. ; KLAUER, S. ; NEALE, V. ; PETERSEN, A. ; LEE, S. ; SUDWEEKS, J. ; PEREZ, M. ; HANKEY, J. ; RAMSEY, D. ; GUPTA, S. ; BUCHER, C. ; DOERZAPH, Z. R. ; JERME-LAND, J. ; KNIPLING, R. : The 100-Car Naturalistic Driving Study, Phase II - Results of the 100 Car Field Experiment. Washington, DC., April 2006 (DOT HS 810 593). – Forschungsbericht

- [29] EBERLY, D. : *Intersection of Ellipsoids*. <http://www.geometrictools.com/Documentation/IntersectionOfEllipsoids.pdf>. Version: 2008
- [30] EL-HAKIM, S. F.: Real-Time Image Meteorology with CCD Cameras. In: *Photogrammetric Engineering and Remote Sensing* 52 (1986), Nr. 11, S. 1757–1766
- [31] FANG, Y. ; MASAKI, I. ; HORN, B. : Depth-Based Target Segmentation for Intelligent Vehicles: Fusion of Radar and Binocular Stereo. In: *IEEE Transaction on Intelligent Transportation Systems* 3 (2002), September, Nr. 3, S. 196–202
- [32] FERNANDO, G. ; CERRI, P. ; BROGGI, A. : Vehicle Detection based on Laser Radar. In: *Intelligent Transportation Systems*, 2007, S. 95–105
- [33] FISCHLER, M. A. ; BOLLES, R. C.: Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography / Communications of the ACM. 1981. – Forschungsbericht
- [34] FITZGIBBON, A. ; PILU, M. ; R.B.FISHER: Direct Least Square Fitting of Ellipses. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* Bd. 21, 1999, S. 476–480
- [35] FORBES, A. B.: Least Squares Best-Fit Geometric Elements. In: *National Physical Laboratory, Report DITC 140/89*, 1989
- [36] FRANKE, U. ; HEINRICH, S. : Fast obstacle detection for urban traffic situations. In: *IEEE Transaction on Intelligent Transportation Systems* 3 (2002), September, Nr. 3, S. 173–181. <http://dx.doi.org/10.1109/TITS.2002.802934>. – DOI 10.1109/TITS.2002.802934
- [37] FRANKE, U. ; ZELL, R. A.: Erkennungsdienst für Gefahren. In: *Daimler High-Tech-Report 1/2009*, Daimler AG, 2009, S. 24–30
- [38] FRANKEN, D. ; HUPPER, A. : Improved fast covariance intersection for distributed data fusion. In: *Proc. 8th Int Information Fusion Conf* Bd. 1, 2005
- [39] FÜRSTENBERG, K. ; DIETMAYER, K. : Fahrzeugumfelderfassung mit mehrzeiligen Laserscannern. In: *Technisches Messen* 3 (2004), S. 164–172
- [40] FÜRSTENBERG, K. C.: *Fahrzeugumfelderfassung und Fußgängerschutz unter Nutzung mehrzeiliger Laserscanner*, Universität Ulm, Diss., 2009
- [41] FÜRSTENBERG, K. C. ; DIETMAYER, K. : Pedestrian Recognition and Tracking of Vehicles using a vehicle based Multilayer Laserscanner. In: *10th World Congress on Intelligent Transport Systems*, 2003
- [42] FÜRSTENBERG, K. : InterSafe2 Final Report / INTERSAFE-2 Consortium. 2011 (D1.2). – Deliverable
- [43] GAN, Q. ; HARRIS, C. J.: Comparison of two measurement fusion methods for Kalman-filter-based multisensor data fusion. 37 (2001), Nr. 1, S. 273–279. <http://dx.doi.org/10.1109/7.913685>. – DOI 10.1109/7.913685

- [44] GANDER, W. ; GOLUB, G. ; STREBEL, R. : Least Square Fitting of Circles and Ellipses. In: *BIT*, 1994, S. 558–578
- [45] GAO, Y. ; RAN, C. ; DENG, Z. : Weighted measurement fusion Kalman filter with correlated measurement noise and its global optimality. In: *Proceedings of the International colloquiumen information Fusion*, 2007
- [46] GARCIA, F. ; JIMEENEZ, F. ; NARANJO, J. E. ; ZATO, J. G. ; APARICIO, F. ; ARMINGOL, J. M. ; ESCALERA, A. de l.: Analysis of LIDAR sensors for new ADAS applications. Usability in moving obstacles detection. In: *ITS World Congress*. Stockholm, 2009
- [47] GORDON, B. : *Zur Bestimmung von Messunsicherheiten terrestrischer Laserscanner*, Technischen Universität Darmstadt, Diss., 2008
- [48] GORJESTANI, A. ; MENON, A. ; CHENG, P.-M. ; SHANKWITZ, C. ; DONATH, M. : The Design of a Minimal Sensor Configuration for a Cooperative Intersection Collision Avoidance System - Stop Sign Assist / Intelligent Vehicles Laboratory, Department of Mechanical Engineering, University of Minnesota. 2010. – Forschungsbericht
- [49] HABERJAHN, M. : Low-Level-Fusion eines Laserscanner- und Stereokamera-Systems in der Fahrzeugumfelderfassung. In: *Proc. of 3D-Nordost*. Berlin, 9-10 December 2010, S. 35–44
- [50] HABERJAHN, M. ; JUNGHANS, M. : Vehicle Environment Detection by a Combined Low and Mid Level Fusion of a Laser Scanner and Stereo Vision. In: *Proc. 14th International IEEE Conference on Intelligent Transportation Systems*. Washington DC, USA, 5.-7. Oct. 2011
- [51] HABERJAHN, M. ; REULKE, R. : Cross-Kalibrierung eines Mehrzeilen-Laserscanner- und Stereokamera-Systems zur Fahrzeugumfelderfassung. In: *Proc. of 3D-Nordost*. Berlin, 10-11 December 2009, S. 119–128
- [52] HABERJAHN, M. ; REULKE, R. : Vehicle Object Discrimination and Tracking by a combined Laser Scanner Stereo System. In: *The 10th Asian Conference on Computer Vision*, 2010
- [53] *Kapitel 1: Multisensor Data Fusion*. In: HALL, D. L. ; LLINAS, J. : *Handbook of Multisensor Data Fusion: Theory and Practice*. 2. CRC Press, 2009, S. 1–14
- [54] HALL, D. ; MCMULLEN, S. A. H.: *Mathematical Techniques in Multisensor Data Fusion*. 2. Artech House Inc, 2004
- [55] HIRSCHMULLER, H. : Accurate and efficient stereo processing by semi-global matching and mutual information. In: *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition CVPR 2005* Bd. 2, 2005, S. 807–814
- [56] HOEGER, R. : HAVEit Final Report / HAVEit Consortium. 2011 (D61.1). – Deliverable
- [57] HORN, B. K. P.: Closed-form solution of absolute orientation using unit quaternions. In: *Journal of the Optical Society of America* 4 (1987), Nr. 4, S. 629–642

- [58] HUANG, Y. ; FU, S. ; THOMPSON, C. : Stereovision-Based Object Segmentation for Automotive Applications. In: *EURASIP Journal on Applied Signal Processing* 14 (2005), S. 2322–2329. <http://dx.doi.org/1155/ASP.2005.2322>. – DOI 1155/ASP.2005.2322
- [59] IBEO AUTOMOTIVE SYSTEMS GMBH: *Betriebsanleitung ibeo LUX® Laserscanner*. Ibeo Automobile Sensor GmbH, Merckurung 20 D, 22143 Hamburg: Ibeo Automobile Sensor GmbH, 2009
- [60] *Kapitel 14: General Decentralized Data Fusion with Covariance Intersection*. In: JULIER, S. ; UHLMANN, J. K.: *Handbook of Multisensor Data Fusion: Theory and Practice*. 2. CRC Press, 2009, S. 319–344
- [61] KERN, F. : Prüfen und Kalibrieren von terrestrischen Laserscannern. In: *Optische 3D-Messtechnik, Photogrammetrie, Laserscanning*. Heidelberg : Wichmann, 2008, S. 306–316
- [62] KESSLER, O. : Functional Description of the Data Fusion Process / Warminster, PA: Office of Naval Technology, Naval Air Development Center. 1992. – Forschungsbericht
- [63] KIRCHNER, A. ; WEISSER, H. ; SCHARNHORST, T. ; STÜKER, D. : Sensor fusion - a research perspective. In: *Vehicle Safety - New Concepts for Automotive Safety*. München, 2001
- [64] KÄMPCHEN, N. : *Feature-Level Fusion of Laser Scanner and Video Data for Advanced Driver Assistance Systems*, Universität Ulm, Diss., 2007
- [65] KRAUS, K. : *Photogrammetrie, Geometrische Informationen aus Photographien und Laserscanneraufnahmen*. Bd. 1. 7. Auflage. Walter de Gruyter, 2004. – 28–32 S.
- [66] KUMAR, S. ; GUPTA, D. ; YADAV, S. : Sensor Fusion of Laser & Stereo Vision Camera for Depth Estimation and Obstacle Avoidance. In: *International Journal of Computer Applications* 1 (2010), Nr. 26, S. 22–26
- [67] LABAYRADE, R. ; ROYERE, C. ; GRUYER, D. ; AUBERT, D. : Cooperative Fusion for Multi-Obstacles Detection With Use of Stereovision and Laser Scanner. In: *Autonomous Robots* 19 (2005), S. 117–140
- [68] LEE, Y.-C. ; HSIAO, T. : Object tracking via the probability-based segmentation using laser range images. In: *Proc. IEEE Intelligent Vehicles Symp. (IV)*, 2010, S. 197–202
- [69] LIGGINS, I. M. E. M. E. ; CHONG, C.-Y. ; KADAR, I. ; ALFORD, M. G. ; VANNICOLA, V. ; THOMOPOULOS, S. : Distributed fusion architectures and algorithms for target tracking. 85 (1997), Nr. 1, S. 95–107. <http://dx.doi.org/10.1109/JPROC.1997.554211>. – DOI 10.1109/JPROC.1997.554211
- [70] LOWE, D. G.: Object recognition from local scale-invariant features. In: *Proc. Seventh IEEE International Conference on Computer Vision* Bd. 2, 1999, S. 1150–1157
- [71] LUHMANN, T. : *Nahbereichsphotogrammetrie*. 2. Auflage. Wichmann, 2003

- [72] MAHALANOBIS, C. P.: On the generalized distance in statistics. In: *Proceedings of the National Institute of Science*. Calcutta, 1936, S. 49–55
- [73] MANYIKA, J. ; DURRANT-WHYTE, H. : *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. Upper Saddle River, NJ, USA : Prentice Hall PTR, 1995
- [74] MATZKA, S. ; ALTENDORFER, R. : A comparison of track-to-track fusion algorithms for automotive sensor fusion. In: *Proc. IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems MFI 2008*, 2008, S. 189–194
- [75] MAYBECK, P. S.: *Stochastic Models, Estimation, and Control Volume 1*. Academic Press, 1979
- [76] MÄHLISCH, M. ; SZCZOT, M. ; LÖHLEIN, O. ; MUNZ, M. ; DIETMAYER, K. : Simultaneous processing of multitarget state measurements and object individual sensory existence evidence with the joint integrated probabilistic data association filter. In: *Proc. 5th Int. Workshop on Intelligent Transportation*. Hamburg, Germany, March 2008
- [77] MITCHELL, H. B.: *Multi-Sensor Data Fusion: An Introduction*. 1st. Springer Publishing Company, Incorporated, 2007. – ISBN 3540714634, 9783540714637
- [78] MORI, S. ; BARKER, W. H. ; CHONG, C.-Y. ; CHANG, K.-C. : Track association and track fusion with nondeterministic target dynamics. 38 (2002), Nr. 2, S. 659–668. <http://dx.doi.org/10.1109/TAES.2002.1008994>. – DOI 10.1109/TAES.2002.1008994
- [79] MUNZ, M. ; MÄHLISCH, M. ; DIETMAYER, K. : Generic Centralized Multi Sensor Data Fusion Based on Probabilistic Sensor and Environment Models for Driver Assistance Systems. 2 (2010), Nr. 1, S. 6–17. <http://dx.doi.org/10.1109/MITS.2010.937293>. – DOI 10.1109/MITS.2010.937293
- [80] NEDEVSCHI, S. ; TIBERIU, M. ; DANESCU, R. ; ONIGA, F. ; BOTA, S. : On-board stereo sensor for intersection driving assistance architecture and specification. In: *Proc. IEEE 5th Int. Conf. Intelligent Computer Communication and Processing ICCP 2009*, 2009, S. 409–416
- [81] NEDEVSCHI, S. ; DANESCU, R. ; FRENTIU, D. ; MARITA, T. ; ONIGA, F. ; POCOL, C. ; GRAF, T. ; SCHMIDT, R. : High Accuracy Stereovision Approach for Obstacle Detection on Non-Planar Roads. In: *Proceedings IEEE Intelligent Engineering Systems*, 2004, S. 211–216
- [82] NIEMEIER, W. : *Ausgleichungsrechnung*. 2. Auflage. de Gruyter, 2008. – 50–72 S.
- [83] PARK, S.-B. : ProFusion2 Final report / PReVENT Consortium 2006. 2007 (D15.12). – Deliverable
- [84] PERROLLAZ, M. ; LABAYRADE, R. ; ROYERE, C. ; HAUTIERE, N. ; AUBERT, D. : Long Range Obstacle Detection Using Laser Scanner and Stereovision. In: *Proc. IEEE Intelligent Vehicles Symp*, 2006, S. 182–187

- [85] PILU, M. : Uncalibrated Stereo Correspondence by Singular Value Decomposition / HP Laboratories Bristol. 1997 (Technical Report HPL-97-96). – Forschungsbericht
- [86] PREMEBIDA, C. ; MONTEIRO, G. ; NUNES, U. ; PEIXOTO, P. : A Lidar and Vision-based Approach for Pedestrian and Vehicle Detection and Tracking. In: *Proc. IEEE Intelligent Transportation Systems Conf. ITSC 2007*, 2007, S. 1044–1049
- [87] PUNSKA, O. : *Bayesian Approaches to Multi-Sensor Data Fusion*, University of Cambridge, Diplomarbeit, 1999
- [88] RAACK, W. : *Mechanik*. 6. Auflage. Technische Universität Berlin. – 106 S.
- [89] REULKE, R. ; LUBER, A. ; HABERJAHN, M. ; PILTZ, B. : Validierung von mobilen Stereokamerasystemen in einem 3D-Testfeld. In: *3D-NordOst*, 2009, S. 45 –52
- [90] RIETDORF, A. : *Automatisierte Auswertung und Kalibrierung von scannenden Messsystemen mit tachymetrischem Messprinzip*, Technische Universität Berlin, Diss., 2005
- [91] ROECKER, J. A. ; MCGILLEM, C. D.: Comparison of two-sensor tracking methods based on state vector fusion and measurement fusion. 24 (1988), Nr. 4, S. 447–449. <http://dx.doi.org/10.1109/7.7186>. – DOI 10.1109/7.7186
- [92] ROSIN, P. : A Note on the Least Squares Fitting of Ellipses. In: *Pattern Recognition Letters*, 1993, S. 799–808
- [93] SANTOLARIA, J. ; PASTOR, J. ; BROSED, F. ; AGUILAR, J. : A one-step intrinsic and extrinsic calibration method for laser line scanner operation in coordinate measuring machines. In: *electronic journals - Measurement Science and Technology* Bd. 20, 2009
- [94] SCARAMUZZA, D. ; HARATI, A. ; SIEGWART, R. : Extrinsic self calibration of a camera and a 3D laser range finder from natural scenes. In: *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems IROS 2007*, 2007, S. 4164–4169
- [95] SCHARSTEIN, D. ; SZELISKI, R. ; ZABIH, R. : A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In: *Proc. IEEE Workshop Stereo and Multi-Baseline Vision (SMBV 2001)*, 2001, S. 131–140
- [96] SCHEUNERT, U. ; LINDNER, P. ; RICHTER, E. ; TATSCHKE, T. ; SCHESTAUBER, D. ; FUCHS, E. ; WANIELIK, G. : Early and Multi Level Fusion for Reliable Automotive Safety Systems. In: *Proc. IEEE Intelligent Vehicles Symp*, 2007, S. 196–201
- [97] SCHMIDT, H. : Dekomposition der endlichen Mischung von Normalverteilungen. In: *Festschrift zur Emeritierung von Univ.-Prof. Dr.-Ing. Bertold Witte*, 2002, S. 131 – 143
- [98] SCOTT, G. ; LONGUET-HIGGINS, H. : An algorithm for associating the features of two patterns. In: *Proc. Royal Society London* Bd. B244, 1991, S. 21–26
- [99] SHAFER, G. : *A Mathematical Theory of Evidence*. Princeton University Press, 1976

- [100] SMITH, K. ; GATICA-PEREZ, D. ; ODOBEZ, J.-M. ; BA, S. : Evaluating Multi-Object Tracking. In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Empirical Evaluation Methods in Computer Vision*. San Diego, 2005
- [101] SPANGENBERG, R. ; DÖRING, T. : Evaluation of object tracking in traffic scenes. In: *Proceedings of the ISPRS Commission V Symposium on Image Engineering and Vision Metrology*. Dresden, Germany, 2006
- [102] SPINELLO, L. ; SIEGWART, R. : Human detection using multimodal and multidimensional features. In: *Proc. IEEE Int. Conf. Robotics and Automation ICRA 2008*, 2008, S. 3264–3269
- [103] STATISTISCHES BUNDESAMT: *Unfallentwicklung auf deutschen Strassen 2010*. Wiesbaden: Statistisches Bundesamt, 2011
- [104] STATISTISCHES BUNDESAMT: *Verkehr - Verkehrsunfälle*. Fachserie 8 Reihe 7. Wiesbaden: Statistisches Bundesamt, 2011
- [105] STRAUBE, M. : Volkswirtschaftliche Kosten durch Straßenverkehrsunfälle in Deutschland 2008. In: *Berichte der Bundesanstalt für Straßenwesen, Unterreihe - Mensch und Sicherheit* (2010), Nr. 17. http://www.bast.de/cln_007/nn_622184/SharedDocs/Publikationen/Forschung-kompakt/2010-17,templateId=raw,property=publicationFile.pdf/2010-17.pdf
- [106] SUESSE, H. ; VOSS, K. : A new efficient algorithm for fitting of rectangles and squares. In: *Proc. Int Image Processing Conf Bd. 2*, 2001, S. 809–812
- [107] TRIMBLE NAVIGATION LIMITED: *DATENBLATT: Trimble 5700 GPS-EMPFÄNGER*. California: Trimble Navigation Limited, 2007. http://trl.trimble.com/dscgi/ds.py/Get/File-171247/022543-074E-D_5700_DS_0808_LR.pdf
- [108] UHLMANN, J. K.: General Data Fusion for Estimates with Unknown Cross Covariances. In: *SPIE* 2755 (1996), S. 536–547
- [109] UHLMANN, J. K.: Covariance consistency methods for fault-tolerant distributed data fusion. In: *Information Fusion* 4 (2003), Nr. 3, S. 201–215
- [110] UMNICKRISHNAN, R. ; HEBERT, M. : Fast Extrinsic Calibration of a Laser Rangefinder to a Camera / Robotics Institute, Carnegie Mellon University. 2005. – Forschungsbericht
- [111] WANG, Y. ; LI, X. R.: A fast and fault-tolerant convex combination fusion algorithm under unknown cross-correlation. In: *Proc. 12th Int. Conf. Information Fusion FUSION '09*, 2009, S. 571–578
- [112] WELCH, G. ; BISHOP, G. : An Introduction to the Kalman Filter / University of North Carolina at Chapel Hill. 1995. – Forschungsbericht
- [113] WENDER, S. ; DIETMAYER, K. : 3D Vehicle Detection using a Laser Scanner and a Video Camera. In: *6th European Congress on ITS*. Aalborg, 2007

- [114] WENDER, S. ; FUERSTENBERG, K. C. ; DIETMAYER, K. : Object Tracking and Classification for Intersection Scenarios Using A Multilayer Laserscanner. In: *Proceedings of ITS 2004*, 2004
- [115] YERKES, R. ; DODSON, J. : The relation of strength of stimulus to rapidity of habit-formation. In: *Journal of Comparative Neurology and Psychology* 18 (1908), S. 459–482
- [116] ZHU, J. ; WANG, L. ; YANG, R. ; DAVIS, J. E. ; PAN, Z. : Reliability Fusion of Time-of-Flight Depth and Stereo Geometry for High Quality Depth Maps. 33 (2011), Nr. 7, S. 1400–1414. <http://dx.doi.org/10.1109/TPAMI.2010.172>. – DOI 10.1109/TPAMI.2010.172
- [117] ZITNICK, C. L. ; KANADE, T. : A cooperative algorithm for stereo matching and occlusion detection. 22 (2000), Nr. 7, S. 675–684. <http://dx.doi.org/10.1109/34.865184>. – DOI 10.1109/34.865184

Anhang A

1 Transformation der Referenzpunkte

Die Gleichungen 1 beschreiben die Umrechnung einer Referenzpunktposition R' in jede andere Referenzpunktposition R'' . Durch Umstellung der Gleichungen lassen sich die fehlenden Paarungen ableiten.

$$\begin{aligned}
 x_{R1} &= x_{R2} + \sin(\psi_O)d_y, & y_{R1} &= y_{R2} - \cos(\psi_O)d_y \\
 x_{R1} &= x_{R3} + \cos(\psi_O)d_x, & y_{R1} &= y_{R3} + \sin(\psi_O)d_x \\
 x_{R1} &= x_{R4} + \cos(\psi_O)d_x + \sin(\psi_O)d_y, & y_{R1} &= y_{R4} + \sin(\psi_O)d_x - \cos(\psi_O)d_y \\
 x_{R1} &= x_C + \cos(\psi_O)\frac{d_x}{2} + \sin(\psi_O)\frac{d_y}{2}, & y_{R1} &= y_C + \sin(\psi_O)\frac{d_x}{2} - \cos(\psi_O)\frac{d_y}{2} \\
 x_{R2} &= x_{R3} + \cos(\psi_O)d_x - \sin(\psi_O)d_y, & y_{R2} &= y_{R3} + \sin(\psi_O)d_x + \cos(\psi_O)d_y \\
 x_{R2} &= x_{R4} + \cos(\psi_O)d_x, & y_{R2} &= y_{R4} + \sin(\psi_O)d_x \\
 x_{R2} &= x_C + \cos(\psi_O)\frac{d_x}{2} - \sin(\psi_O)\frac{d_y}{2}, & y_{R2} &= y_C + \sin(\psi_O)\frac{d_x}{2} + \cos(\psi_O)\frac{d_y}{2} \\
 x_{R3} &= x_{R4} + \sin(\psi_O)d_y, & y_{R3} &= y_{R4} - \cos(\psi_O)d_y \\
 x_{R3} &= x_C - \cos(\psi_O)\frac{d_x}{2} + \sin(\psi_O)\frac{d_y}{2}, & y_{R3} &= y_C - \sin(\psi_O)\frac{d_x}{2} - \cos(\psi_O)\frac{d_y}{2} \\
 x_{R4} &= x_C - \cos(\psi_O)\frac{d_x}{2} - \sin(\psi_O)\frac{d_y}{2}, & y_{R4} &= y_C - \sin(\psi_O)\frac{d_x}{2} + \cos(\psi_O)\frac{d_y}{2}
 \end{aligned} \tag{1}$$

Die restlichen Größen $z_{R'}$, $d_{x_{R'}}$, $d_{y_{R'}}$, $d_{z_{R'}}$ und $\omega_{R'}$ des vorherigen Objektzustands \mathbf{X}' sowie ihre Genauigkeitsangaben werden ohne Transformation in den neuen Zustand \mathbf{X}'' übernommen:

$$\begin{aligned}
 z_{R''} &= z_{R'}, & d_{x_{R''}} &= d_{x_{R'}}, & d_{y_{R''}} &= d_{y_{R'}}, \\
 d_{z_{R''}} &= d_{z_{R'}}, & \omega_{R''} &= \omega_{R'}
 \end{aligned} \tag{2}$$

Die hieraus abzuleitenden Differentialquotienten zum Aufstellen der Jacobi-Matrix \mathbf{A} und zur Berechnung der neuen Kovarianzmatrix \mathbf{P}'' sind aufgrund ihrer Einfachheit hier nicht aufgeführt. Die Varianzen und die Kovarianzen zu den Zustandsgrößen ψ_O und ω_O werden in \mathbf{P}'' aus \mathbf{P}' übernommen.

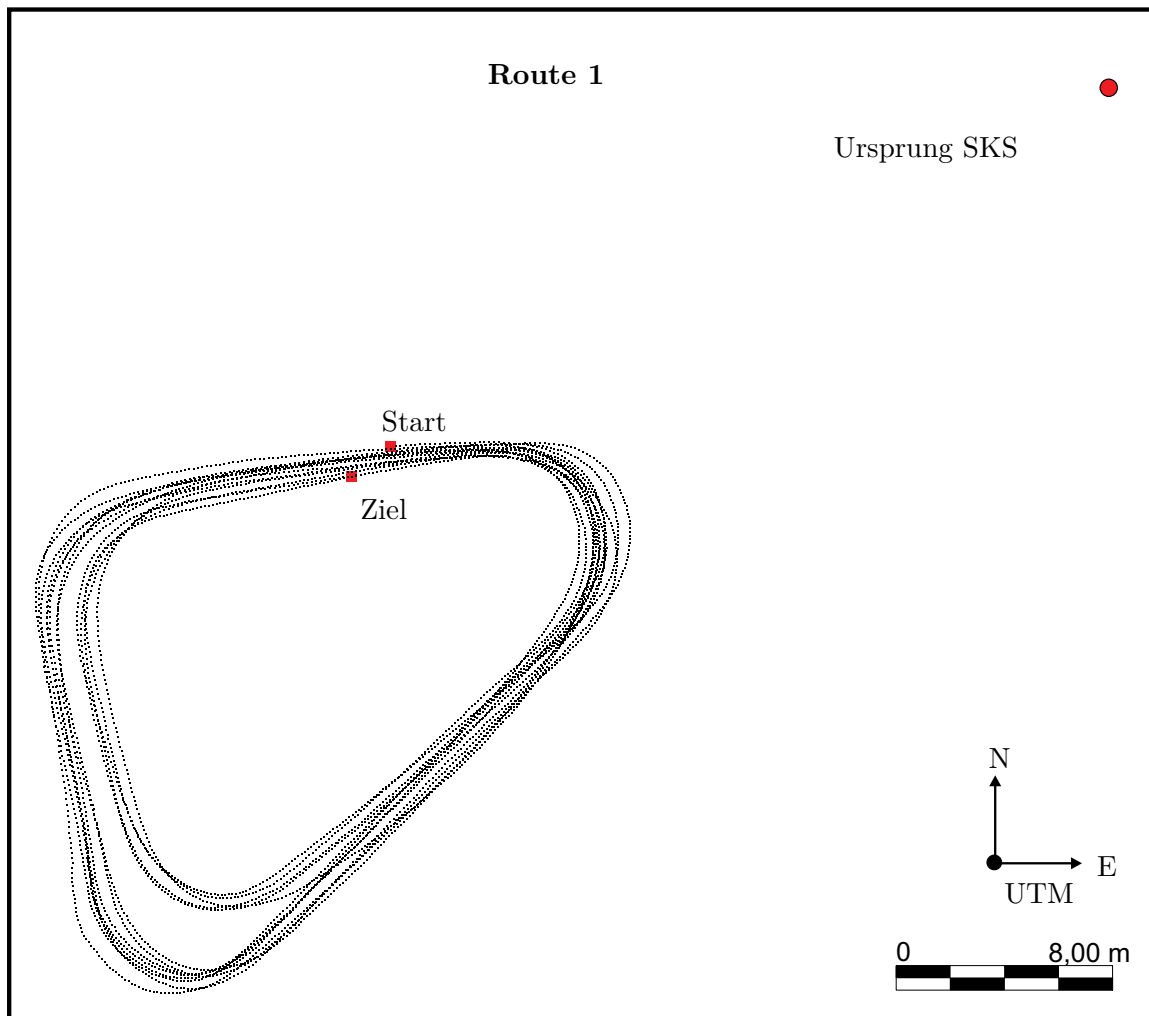


Abbildung 1: Route 1 gemessen an der hinteren Antennenposition am Referenzfahrzeug

2 Parameter zur Experimentauswertung

Die drei Messrouten zur Bestimmung der Objektbox-Genauigkeit sind in den Abbildungen 1, 2 und 3 dargestellt. Die einzelnen Messpositionen stammen hierbei von der hinteren GPS-Antenne des Referenzfahrzeugs (siehe Abschnitt 7.6).

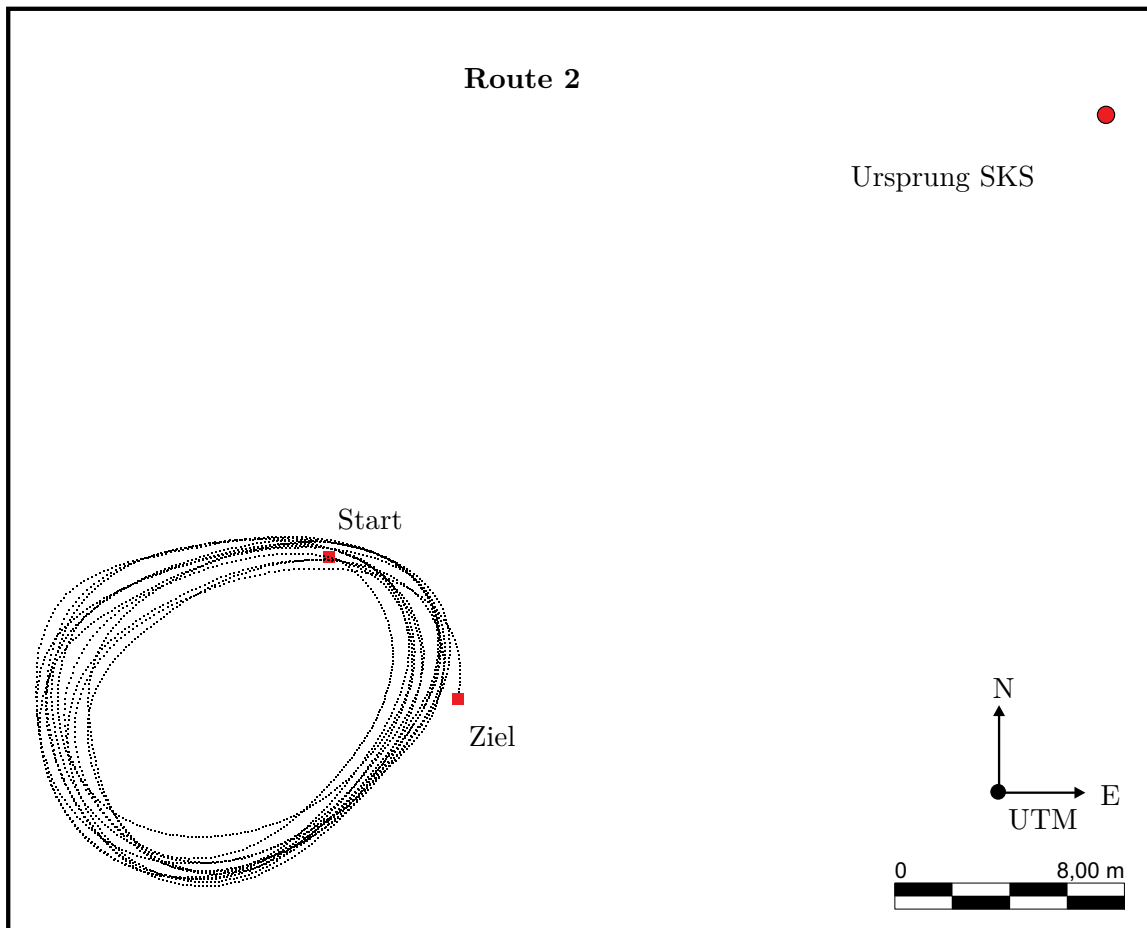


Abbildung 2: Route 2 gemessen an der hinteren Antennenposition am Referenzfahrzeug

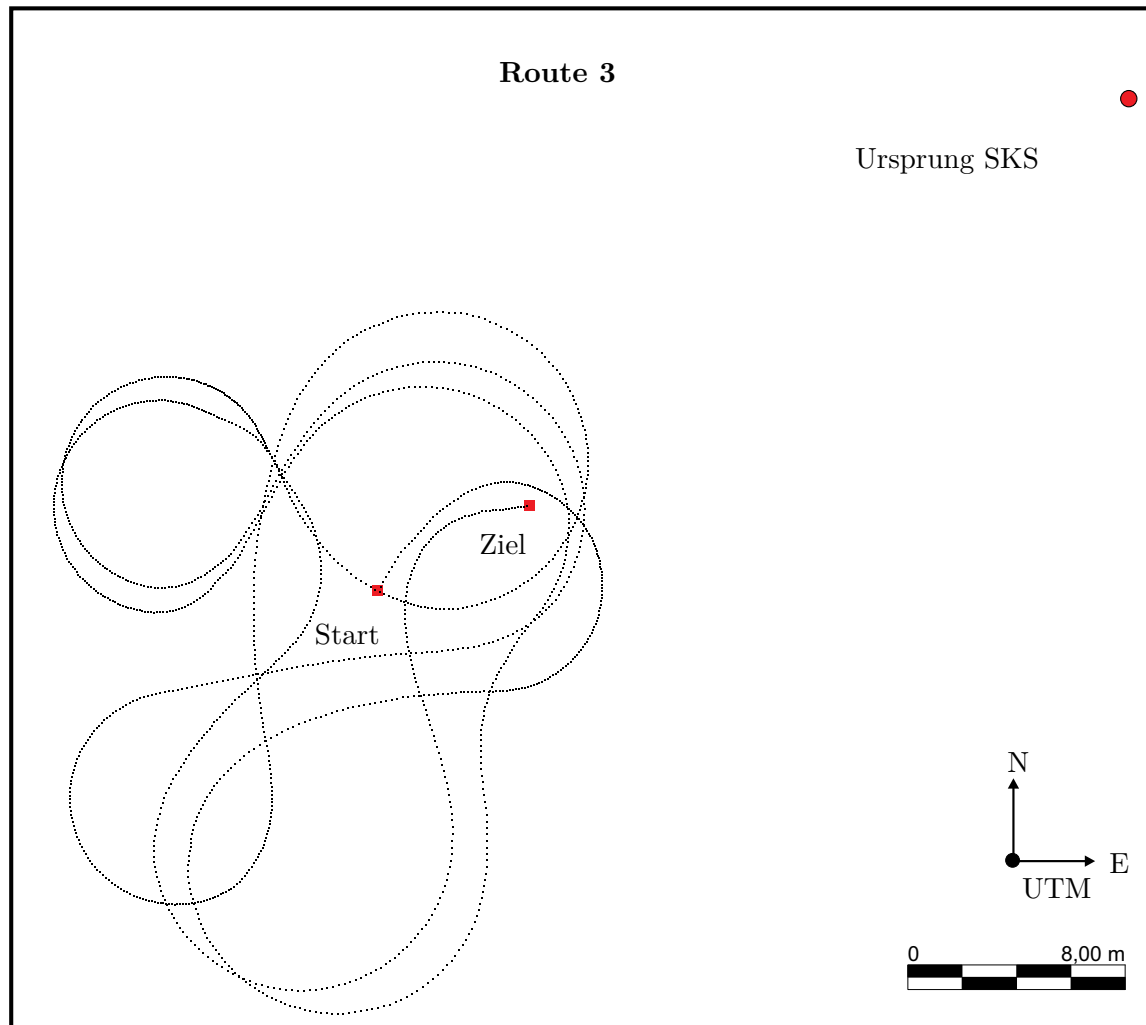


Abbildung 3: Route 3 gemessen an der hinteren Antennenposition am Referenzfahrzeug

Rotationsmatrix

Bildung der Rotationsmatrix R anhand der Euler-Winkel ω , ϕ und κ :

$$\begin{aligned}
 R &= R_x R_y R_z \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & -\sin\omega \\ 0 & \sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\kappa & -\sin\kappa & 0 \\ \sin\kappa & \cos\kappa & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)
 \end{aligned}$$

Parameter Stereokamera

Innere Orientierung linke Kamera

Parameter	Wert	σ
c in mm	12.4509	$2.78 \cdot 10^{-4}$
x_p in mm	0.2243	$4.29 \cdot 10^{-4}$
y_p in mm	0.0199	$4.2 \cdot 10^{-4}$
k_1	$3.79 \cdot 10^{-4}$	$2.11 \cdot 10^{-6}$
k_2	$-1.64 \cdot 10^{-5}$	$1.59 \cdot 10^{-7}$
k_3	$-1.69 \cdot 10^{-7}$	$3.62 \cdot 10^{-9}$
p_1	$5.88 \cdot 10^{-5}$	$1.05 \cdot 10^{-6}$
p_2	$-5.22 \cdot 10^{-5}$	$9.85 \cdot 10^{-7}$
b_1	$1.11 \cdot 10^{-4}$	$2.99 \cdot 10^{-6}$
b_2	$-1.09 \cdot 10^{-5}$	$2.91 \cdot 10^{-64}$

Tabelle 1: Parameter der inneren Orientierung der linken Kamera des Stereokamerasystems (siehe Abschnitt 3.2.1)

Innere Orientierung rechte Kamera

Parameter	Wert	σ
c in mm	12.4846	$3.29 \cdot 10^{-4}$
x_p in mm	-0.0309	$5.19 \cdot 10^{-4}$
y_p in mm	0.0629	$5.01 \cdot 10^{-4}$
k_1	$2.49 \cdot 10^{-4}$	$2.77 \cdot 10^{-6}$
k_2	$-9.54 \cdot 10^{-6}$	$2.23 \cdot 10^{-7}$
k_3	$-2.87 \cdot 10^{-7}$	$5.37 \cdot 10^{-9}$
p_1	$1.21 \cdot 10^{-4}$	$1.26 \cdot 10^{-6}$
p_2	$1.24 \cdot 10^{-4}$	$1.16 \cdot 10^{-6}$
b_1	$-2.31 \cdot 10^{-5}$	$3.77 \cdot 10^{-6}$
b_2	$-4.31 \cdot 10^{-6}$	$3.64 \cdot 10^{-6}$

Tabelle 2: Parameter der inneren Orientierung der rechten Kamera des Stereokamerasystems (siehe Abschnitt 3.2.1)

Relative Orientierung

Parameter	Wert
T_x	0
T_y	0
T_z	1
ω_L in rad	$5.79 \cdot 10^{-3}$
ϕ_L in rad	$-1.43 \cdot 10^{-4}$
κ_L in rad	$-7.99 \cdot 10^{-3}$
ω_R in rad	0
ϕ_R in rad	$3.12 \cdot 10^{-3}$
κ_R in rad	$-3.42 \cdot 10^{-3}$
RMS	$-2.14 \cdot 10^{-3}$

Tabelle 3: Parameter der relativen Orientierung nach Bilddrehung der Stereokamera (siehe Abschnitt 3.2.1)

Absolute Orientierung

Parameter	Wert
T_x in mm	3875.61
T_y in mm	2799.29
T_z in mm	1725.91
ω in rad	-1.31953
ϕ in rad	0.97765
κ in rad	2.91809
m_a	399.00692
RMS in mm	0.7073

Tabelle 4: Parameter der absoluten Orientierung der Stereokamera zwischen Modellkoordinatensystem der relativen Orientierung und Weltkoordinatensystem der Referenzkoordinaten im Labor (siehe Abschnitt 3.2.1)

Segmentierungsmodell

Parameter	Wert
O_x in mm	50
O_y in mm	20
O_z in mm	0
S_x	0.025
S_y	0.025
S_z	0.025
E_x	2
E_y	1
E_z	1
Messauflösung horizontal in rad	0.00052
Messauflösung vertikal in rad	0.00052

Tabelle 5: Parameter des Segmentierungsmodells der Stereokamera (siehe Gleichung 4.1)

Stereo Matcher

Parameter	Wert
<i>SAD window size</i> in Pixel	13
<i>Number of disparities</i> in Pixel	48
<i>min disparity</i> in Pixel	1
<i>Uniqueness ratio</i>	20
<i>Texture threshold</i>	20
<i>Pre filter cap</i>	38
<i>Pre filter size</i>	21
<i>Speckle range</i>	10
<i>Speckle window size</i>	19

Tabelle 6: Parameter des *openCV* Block Matcher zur Stereokorrespondenz-Analyse (siehe Abschnitt 3.1.1)

Parameter Laserscanner

Fehlermodell

Parameter	Wert
a	1.0036
b in mm	52.0872
s	1.0541
σ_r in mm	31.6711
σ_φ in rad	$1 \cdot 10^{-5}$
σ_θ in rad	$3 \cdot 10^{-4}$
RMS in mm	14.8809

Tabelle 7: Parameter des Fehlermodells des Laserscanners (siehe Gleichungen 3.10, 3.11, 3.19 und 3.23)

Äußere Orientierung zu SKS

Parameter	Wert
T_x in mm	-142.5331
T_y in mm	-201.471
T_z in mm	2.3364
ω in rad	$-7.58 \cdot 10^{-3}$
ϕ in rad	$-5.37 \cdot 10^{-3}$
κ in rad	$-3.39 \cdot 10^{-3}$
RMS in mm	14.8809

Tabelle 8: Parameter der äußeren Orientierung des Laserscanners LKS zum SKS (siehe Gleichung 3.17)

Segmentierungsmodell

Parameter	Wert
O_x in mm	1
O_y in mm	20
O_z in mm	1
S_x	0.035
S_y	0.02
S_z	0.04
E_x	1
E_y	1
E_z	1
Messauflösung horizontal in rad	0.00436
Messauflösung vertikal in rad	0.01396

Tabelle 9: Parameter des Segmentierungsmodells des Laserscanners (siehe Gleichung 4.1)

Parameter EKF

Parameter	Wert
σ_x^2 in m^2	1
σ_y^2 in m^2	1
σ_z^2 in m^2	1
$\sigma_{d_x}^2$ in m^2	2
$\sigma_{d_y}^2$ in m^2	2
$\sigma_{d_z}^2$ in m^2	2
σ_v^2 in $(m/ms)^2$	$1 \cdot 10^{-4}$
$\sigma_{\psi_R}^2$ in rad^2	1
$\sigma_{\omega_R}^2$ in $(rad/ms)^2$	$5 \cdot 10^{-8}$
$\sigma_{\psi_O}^2$ in rad^2	1
$\sigma_{\omega_O}^2$ in $(rad/ms)^2$	$5 \cdot 10^{-8}$

Tabelle 10: Initialwerte für die Kovarianzen der Matrix P des EKF (siehe Gleichung 5.13)

Parameter	Wert
$\sigma_{a_v}^2$ in m	$1 \cdot 10^{-14}$
$\sigma_{a_\omega}^2$ in m	$1 \cdot 10^{-14}$
σ_d^2 in m	$1 \cdot 10^{-10}$

Tabelle 11: Initialwerte für die Kovarianzen der Matrix Q des EKF (siehe Gleichung 5.18)

Danksagung

Mein Dank gilt in erster Linie meinem Doktorvater Herr Professor Ralf Reulke für die fachliche Unterstützung während der Promotionszeit.

Ich danke zudem meiner Arbeitsgruppe Verkehrsdatenerfassung am Deutschen Zentrum für Luft- und Raumfahrt in Berlin, insbesondere Marek Junghans, für die vielen inspirierenden Diskussionen und Ratschläge.

Meinen Eltern verdanke ich neben einer günstigen Veranlagung eine ausgezeichnete Ausbildung sowie die nötige Portion Ehrgeiz.

Ganz besonders danke ich jedoch meiner Frau Maria und ihrer Familie. Maria hat mir stets Mut zugesprochen und mich mit Unterstützung und Verständnis beschenkt. Hätte sie mir nicht den Rücken freigehalten, wäre meine Arbeit in dieser Form nicht möglich gewesen.